Towards Learning Analytics Scalability:
Enriching Context Descriptions Across Data Sources

Jeanette Samuelsen

Thesis for the degree of Philosophiae Doctor (PhD)
At the University of Bergen

2023

Scientific environment

The research presented in this dissertation was conducted in the scientific milieu at The Centre for the Science of Learning & Technology (SLATE), the national centre for learning analytics. SLATE is funded by the Ministry of Education and the University of Bergen. SLATE is an interdisciplinary centre of researchers with backgrounds in information science, pedagogy, sociology, informatics, psychology, music, fine arts and law.

During my PhD research, I have been employed by the University of Bergen. First, at the Department of Information Science and Media Studies, and subsequently at SLATE. This research has been supported by a University of Bergen PhD stipend at SLATE and funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 856954.

Acknowledgements

It is quite a special feeling to finally submit the dissertation after all these years. This thesis could not have been completed without all the people that provided their support and input.

First and foremost, I would like to thank my main supervisor Professor Barbara Wasson and my co-supervisor Professor Weiqin Chen.
Barbara: Thank you for all your guidance, enthusiasm, and encouragement. Thank you also for great leadership in your role as the director of SLATE.
Weiqin: Thank you for all your help, patience, and for always being there when I was ready to take the next step. I am also grateful that you introduced me to the Learning Analytics field and research in general.

I would like to extend my thank you to current and former colleagues at SLATE. In particular, I would like to thank the SLATE leadership that in addition to director Professor Barbara Wasson also consists of Dr. Ingunn Johanne Ness and Jorunn Viken. Thank you Ingunn for our conversations about life and academia, and Jorunn for your positive attitude and pragmatic approach. When I doubted myself, it really helped me that the three of you expressed confidence that I will finish this PhD. Thank you also to Nina Morlandstø and Cecilie Johanne Slokvik Hansen for kindly and enthusiastically introducing me to the AVT project and its various stakeholders when I was still new to this research. In addition, I wish to thank Hanne Widnes Gravermoen, PhD Coordinator at the Faculty of Social Sciences, for support and encouragement in the end stage of the PhD.

I am thankful to the participants that contributed to the PhD papers through the user testing and interviews (and who unfortunately can not be named for privacy reasons). Thank you for setting aside time to share from your wealth of knowledge. I have learned so much from you.

In addition, I am grateful to the Standards Norway SN/K 186 community, including its leadership and members, for providing a venue where standardization related to EdTech and Learning Analytics is explored in-depth in the Norwegian context.

Thank you to my significant other, Erik Gjertsen, for continued support of my studies, for ensuring I could always have coffee in the morning, and for being exceptionally patient during this last leg of the PhD when I basically had no free time :) Thank you also for all our discussions on technical topics and for providing input on the user test materials. In addition, you took great care of our dog and cat (Set and Naomi) that should also be acknowledged for making life so much better.

I am also grateful to family and friends at home, for showing a bit of interest in my life in academia, but not asking too much about the PhD progress. Especially Beatriz Mendoza, for being an exceptional hiking friend, for all our meaningful conversations, and for supporting me through the ups and downs of doctoral studies and life in general.

Last but not least, doing a PhD can be a lonely experience, but I was fortunate to have great company: Marianne Lorenz (Tübingen, Germany), Vincy Huang (Delft, Netherlands), Deborah Roberts (Brisbane, Australia), Elia Turner (Tel Aviv, Israel), and Mohammad Awadhalla (Abu Dhabi, United Arab Emirates). Thank you for being there during the last 12 months of my PhD studies, as the third paper and thesis progressed (mostly) 50 minutes at a time. And thank you also for support and encouragement, nice and honest conversations, and for providing examples of hard work and determination.

Bergen, April 2023
Jeanette Samuelsen

Abstract

Following advances in technology, digital tools and systems related to education are increasingly being adopted by educational institutions, including learning management systems, student information systems, exam systems, and adaptive learning tools. Such tools and systems generate and store vast amounts of data pertaining to learning. Learning Analytics (LA) is a field concerned with data collection, analysis and reporting regarding learners and their context, aimed at providing insights and improvements related to learning and learning environments. Data integration, the collection and combination of data originating from multiple data sources, can contribute to more precise and useful LA results based on increased data diversity and scale. However, data originating from different sources are often stored in different formats, with varying levels of structure, using different storage technologies, resulting in silos where data are prevented from being integrated and analyzed across sources. As such, data integration has been identified as a key challenge in LA. Data integration is an important aspect of LA scalability, and is tightly linked to interoperability. To support data integration, the use of a learning activity data standard such as xAPI enables the representation of learner data in the most basic form of *actor verb object*, i.e., a learner interacting with a learning object (in a learning environment), where further syntactic structures are included that allow for the registration of learning context data.

This research aims to gain insight into and address challenges for LA data integration, interoperability, and consequently scalability. More specifically, it is focused on identifying gaps in xAPI expressibility and providing solutions to these gaps.

Following a design science research methodology, the research carried out a systematic review on the state of the art related to LA data integration, and a case study. The case study used the the Activity Data for Assessment and Adaptivity (AVT) project that utilizes xAPI for describing K-12 learner data from multiple data sources as the case, and comprised two empirical studies which included interviews and user testing with stakeholders.

The research resulted in three contributions. The first contribution is the identification of a number of challenges related to core issues of LA scalability, through a systematic literature review. The challenges include that learning activity data standards are not widely used, that publications that report on LA research tend to lack details about integration of data in the implemented systems, and a lack of stakeholder involvement (e.g., from teachers and vendors). The second contribution constitutes the 1) identification of xAPI gaps related to consistent data descriptions across sources, focusing on learning context, and 2) development of a conceptual solution in the form of recommendations to promote consistent data descriptions across data sources. Several gaps identified point to a lack of clarity in how to describe xAPI context data, and consequently the recommendations include a unified and hierarchical structuring of xAPI context. The third contribution is a technical solution, implemented based on the recommendations for consistent data descriptions across data sources, and taking into account the K-12 school adaptivity use case. The technical solution was realized through the creation of an xAPI profile, the school adaptivity profile, and evaluated by means of user testing with technical experts, according to usability criteria of usefulness, effectiveness, and satisfaction. The evaluation indicated that the criteria of usefulness and effectiveness were met. Satisfaction was signaled regarding some aspects of the technical solution, while there were some suggestions for improvement.

This research has shown that consistent data descriptions across sources may advance data integration and interoperability, contributing to improving LA scalability. Aspects of the technical solution, including the unified and hierarchical structuring of context descriptions, should be further explored by the xAPI community for inclusion in the standard as a means to promote scalable LA.

Sammendrag

Den teknologiske utviklingen har ført til at utdanningsinstitusjoner i økende grad tar i bruk digitale verktøy og systemer knyttet til utdanning, blant annet læringsplattformer (learning management systems), studentinformasjonssystemer, eksamenssystemer og adaptive læringsverktøy. Slike verktøy og systemer genererer og lagrer store mengder data knyttet til læring. Læringsanalyse er et felt som er opptatt av datainnsamling, analyse og rapportering om lærende og deres kontekst, med sikte på å gi innsikt og forbedringer knyttet til læring og læringsmiljøer. Dataintegrasjon, det vil si innsamling og kombinasjon av data fra flere datakilder, kan bidra til mer presise og nyttige læringsanalyse resultater basert på økt mangfold (diversity) og omfang (scale) av data. Data fra ulike kilder lagres imidlertid ofte i ulike formater, med varierende struktur og ved hjelp av ulike lagringsteknologier, noe som resulterer i siloer der data forhindres fra å bli integrert og analysert på tvers av kilder. Dataintegrasjon har derfor blitt identifisert som en sentral utfordring i læringsanalyse. Dataintegrasjon er et viktig aspekt ved skalerbarhet (scalability) knyttet til læringsanalyse, og er tett knyttet til interoperabilitet. For å støtte dataintegrasjon, muliggjør bruk av en datastandard for læringsaktiviteter (learning activity data standard), som xAPI, representasjon av data om lærende i den mest grunnleggende formen *aktør-verb-objekt*, dvs. en lærende (for eksempel en elev) som samhandler med et læringsobjekt (i et læringsmiljø), hvor ytterligere syntaktiske strukturer er inkludert som gir mulighet for registrering av læringskontekstdata.

Denne forskningen tar sikte på å få innsikt i og håndtere utfordringer for dataintegrasjon, interoperabilitet og følgelig skalerbarhet knyttet til læringsanalyse. Mer spesifikt er det fokusert på å identifisere mangler i uttrykkbarhet (expressibility) for xAPI, og å bidra med løsninger knyttet til disse manglene.

Forskningen fulgte en overordnet "design science research" forskningsmetodikk, hvor det ble gjennomført en systematisk litteraturgjennomgang knyttet til dataintegrasjon i læringsanalyse, og en casestudie. Casestudien brukte prosjektet "Aktivitetsdata for vurdering og tilpassing" (AVT), som benytter xAPI for å beskrive elevdata fra flere

datakilder, som case. Casestudien omfattet to empiriske studier som inkluderte intervjuer og brukertesting med interessenter.

Forskningen resulterte i tre bidrag. Det første bidraget er identifiseringen av en rekke utfordringer knyttet til sentrale aspekter ved skalerbarhet knyttet til læringsanalyse, gjennom en systematisk litteraturgjennomgang. Utfordringene er blant annet at datastandarder for læringsaktiviteter ikke er mye brukt, at publikasjoner som rapporterer om læringsanalyseforskning har en tendens til å mangle detaljer om integrering av data i de implementerte systemene, og manglende involvering av interessenter (f.eks. lærere og leverandører). Det andre bidraget består av 1) identifisering av xAPI-mangler knyttet til konsistente databeskrivelser på tvers av kilder, med fokus på læringskontekst, og 2) utvikling av en konseptuell løsning i form av anbefalinger for å fremme konsistente databeskrivelser på tvers av datakilder. Flere identifiserte mangler peker på uklarhet i hvordan xAPI-kontekstdata skal beskrives, og anbefalingene inkluderer derfor en enhetlig og hierarkisk strukturering av xAPI-kontekst. Det tredje bidraget er en teknisk løsning som er implementert basert på anbefalingene for konsistente databeskrivelser på tvers av datakilder, og som tar hensyn til scenariet "adaptivitet i grunnopplæringen". Den tekniske løsningen ble realisert gjennom opprettelsen av en xAPI-profil, "school adaptivity profile", og evaluert ved hjelp av brukertesting med tekniske eksperter, i henhold til brukervennlighetskriteriene (usability criteria) "nytte" (utility), "effektivitet" (effectiveness) og "tilfredshet" (satisfaction). Evalueringen viste at brukervennlighetskriteriene nytte og effektivitet ble oppfylt. Det ble signalisert tilfredshet ved noen aspekter ved den tekniske løsningen, men det var også noen forslag til forbedringer.

Denne forskningen har vist at konsistente databeskrivelser på tvers av kilder kan fremme dataintegrasjon og interoperabilitet, og bidra til å forbedre skalerbarhet knyttet til læringsanalyse. Aspekter ved den tekniske løsningen, inkludert enhetlig og hierarkisk strukturering av kontekstbeskrivelser, bør utforskes videre av xAPI fellesskapet (xAPI community) for inkludering i standarden som et middel for å fremme skalerbar læringsanalyse.

List of publications

Paper 1:

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0105-4

Paper 2:

Samuelsen, J., Chen, W., & Wasson, B. (2021). Enriching context descriptions for enhanced LA scalability: a case study. *Research and Practice in Technology Enhanced Learning*, *16*(1). https://doi.org/10.1186/s41039-021-00150-2

Paper 3:

Samuelsen, J., Chen, W., & Wasson, B. (submitted). Implementing enriched context descriptions for efficient scaling of Learning Analytics.
Journal of Learning Analytics.

Contents

List of figures

List of tables

List of abbreviations

ADL	Advanced Distributed Learning
API	Application Programming Interface
AVT	Activity Data for Assessment and Adaptivity
DSR	Design Science Research
ISO	International Organization for Standardization
IEEE	Institute of Electrical and Electronics Engineers
IRI	Internationalized Resource Identifier
JSON-LD	JSON for Linking Data
LA	Learning Analytics
LMS	Learning Management System
LRS	Learning Record Store
RDF	The Resource Description Framework
xAPI	Experience API

Part 1
The extended abstract

1. Introduction

With advances in technology, such as increased computing power and ubiquitous computing, learning resources are increasingly being digitized and digital tools and systems related to education, such as the Learning Management System (LMS), student information system, exam system and adaptive learning tools, are increasingly being adopted by educational institutions and offered online. While the use of technology to enhance and support education is not new, more recent events such as the COVID-19 pandemic and the Ukrainian refugee crisis have particularly underscored the importance of being able to utilize digital technology for education, for instance through the possibility of offering remote learning opportunities.

Within education, the use of digital tools, systems and resources generate a lot of data about the learner[1] that may provide valuable information to instructors, the administration, educational authorities, and not least to the learner themself (Siemens & Long, 2011). Such data are often learning activity data that individually describe the learner interacting with a learning object in a learning environment. How can the activities of a student, group of learners, or institution be better understood? This is a concern for the field of learning analytics (LA), which is commonly defined as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (Siemens, 2011). As this definition clarifies, it is not only data specifically about the learners but also regarding their context that is of importance for LA, where a common definition of context is: "any information that can be used to characterize the situation of an entity" (Dey, 2001, p. 5). Here, the entity may be either a person, a place or an object.

Learner data may be generated by different tools and systems, and the tools and systems may use a variety of storage technologies (e.g., relational database or a variety of NoSQL databases), formats, levels of structure (structured, semi-structured, unstructured), and data retrieval methods; thus, potentially creating a situation where

[1] In this dissertation, the terms learner and student are used interchangeably.

data are isolated in silos and therefore may not be utilized together. This is in contrast to the situation where data from different sources are collected and merged, i.e., the data are integrated. Data integration, however, is a key challenge in learning analytics (Chatti et al., 2017). Integrating data from multiple sources enables richer accounts of individual students' learning activities, thereby contributing to more precise and useful LA results based on increased diversity and scale of the input data (Cooper & Hoel, 2015; Chatti et al., 2017). Related to insights derived from learning analytics, “researchers have suggested that the true potential to offer meaningful insight comes from combining data from across different sources” (Bakharia et al., 2016, p. 378).

Data integration is an important dimension that contributes to LA scalability, where institutional, spatial, and temporal dimensions are also vital (Chatti et al., 2017). Scaling up LA involves not only technical solutions, but additional factors such as organizational and management structures (Buckingham Shum & McKay, 2018; Dawson et al., 2018), in addition to policy and regulation (European Union, 2016).

Interoperability is closely related to data integration. There exist various definitions on interoperability that pertain to ICT and organizations (Cooper, 2013; Benson & Grieve, 2016; European Commission, 2017), where a common characteristic is the emphasis on information exchange and the use of the exchanged information.

Efforts to enhance interoperability have been undertaken for decades. In the late 1980s, Tim-Berners Lee invented the World Wide Web to counter problems where computers were unable to exchange and use the exchanged information. His work addressed the issue of incompatible data formats (Berners-Lee & Fischetti, 1999, p. 35), an issue still relevant today. Looking at technologies in the educational domain, Duval (2004) also emphasized interoperability as a means where information from different sources can be used together.

The adoption of widely used specifications and standards “is crucial for dealing with the data integration [...] and interoperability issues” (Muslim, 2018, p. 30). Two well-known specifications (de facto industry standards) exist related to education and LA, namely the Experience API (xAPI; Advanced Distributed Learning, 2021b)[2] and Caliper Analytics (1EdTech, n.d.-a). These specifications both support the exchange and integration of learning activity data that originate from different tools, services, and data sources, and they both facilitate technical and semantic interoperability. At the most basic level, the data adhering to these specifications are modeled with the constructs of actor, verb, and object[3]. Both specifications have syntactic structures that enable the registration of context data. Furthermore, the specifications both provide a means to specify vocabularies through profiles, where the profiles may also provide rules that may be checked by a profile validator, for example to specify the allowed activity data structure and format for a use case or community of practice.

Despite the potential of xAPI and Caliper in terms of data integration and interoperability, previous research has identified challenges with the standards (e.g., Bakharia et al., 2016; Berg, Mol et al., 2016; Berg, Scheffel et al., 2016). Consequently, it may be of interest to both LA researchers and practitioner communities to understand the challenges and limitations when using these specifications.

Based on this background, the research presented in this dissertation aims to understand and address challenges with regard to LA interoperability, data integration, and consequently scalability. More specifically, the research has focused on identifying gaps and needs for current learning activity data standards, and subsequently on contributing solutions to the identified gaps and needs.

1.1 Research questions and methodology

The overarching research framework for the research is the design science research (DSR) framework (e.g., Hevner et al., 2004; Vaishnavi & Kuechler, n.d.), emphasizing the building and application of innovative artifacts to solve real and important problems. This research framework requires rigorous methods to be used both in building and evaluating the artifacts. Within the DSR framework, the methods of

[2] More precisely, the xAPI standard consists of the xAPI and xAPI profile specifications.
[3] For Caliper Analytics, the terms used are *agent*, *action*, and *object*.

systematic literature review (Kitchenham & Charters, 2007) and case study (Oates, 2006; Bryman, 2012; Yin, 2014) were applied. The case study used a real-world project, the Activity Data for Assessment and Adaptivity[4] (AVT) project, as the case. AVT, a Norwegian research and development project, utilizes the xAPI learning activity data standard for describing K-12 learner data originating from different data sources (EdTech vendors); AVT aims to integrate data to explore areas such as learner adaptivity and assessment.

The three research questions that guided the research included in this dissertation are:

- RQ1: What are the challenges of current LA approaches and practices for data integration, interoperability, and consequently scalability?
- RQ2: How can descriptions of context in a current learning activity data standard, i.e., xAPI, be enhanced in order to provide improved data integration, interoperability, and consequently scalability?
- RQ3: How can the technical solution for enhanced context descriptions in xAPI be validated?

These questions have been addressed in research detailed in the three papers that are part of this dissertation, and which each have their own focus and contributions. The first paper presents a systematic review, while the second and third paper detail the two empirical studies that describe the first and second parts of the case study. An overview of the papers is provided in chapter 4.

Figure 1 gives an overview of the relationship of the research papers resulting from the PhD research.

[4] The Norwegian title of the project is "Aktivitetsdata for vurdering og tilpasning" (AVT).

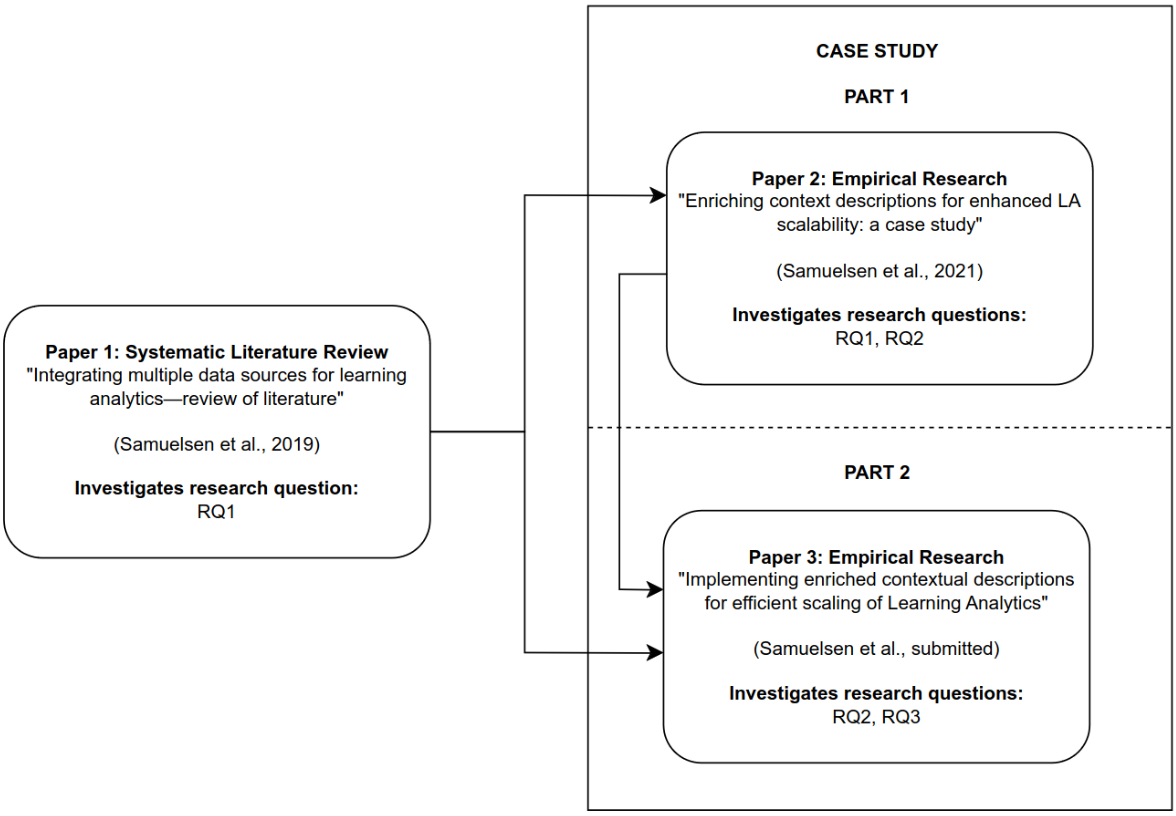

Figure 1 - Overview of the research included in this dissertation

1.2 Contributions

The main contributions from this PhD research to the knowledge base are the following:

C1: Challenges identified related to LA core issues of scalability

This contribution is founded on a systematic review. The results from this review provide an overview of the current state of LA in terms of data integration in the context of higher education. Through gaining insight regarding LA and data integration based on the research literature, a number of shortcomings were identified; addressing the shortcomings may lead to improvements in LA scalability. As such, the review concludes with recommendations to address identified challenges. This contribution is detailed in Paper 1.

C2: Gaps related to xAPI expressibility and conceptual solution

The research identifies gaps and provides a conceptual level solution related to xAPI expressibility - i.e., the consistent description of data across sources, with a focus on learning context descriptions. The gaps were identified through a thematic analysis of interviews with stakeholders from the AVT project, and an inspection of xAPI. A conceptual solution was developed to address identified gaps in the xAPI and xAPI profile specifications related to xAPI expressibility and context structures. This was derived from the xAPI and xAPI profile specifications and a study of previous research on context data models, and resulted in the conceptual solution as a set of recommendations. Among the recommendations is the use of a hierarchical and unified structure to represent xAPI context. The conceptual solution constitutes the first artifact developed as part of this PhD research following the design science research methodology. This contribution is detailed in Paper 2.

C3: Technical solution implemented and validated

A technical solution (instantiation) was developed to implement the recommendations for improved xAPI expressibility, with a particular emphasis on context. The instantiation was implemented based on a K-12 school adaptivity use case, and carefully considered the available functionality and constraints of the xAPI and xAPI profile specifications. The implementation of the technical solution was enabled through creation of an xAPI profile, i.e., the school adaptivity profile. The technical solution was validated through user testing, including complementary interviews with technical experts from the AVT project. The technical solution constitutes the second artifact developed as part of this PhD research following the design science research methodology. The contribution is detailed in Paper 3.

1.3 Organization of the dissertation

This dissertation is organized in two main parts. Part I includes background, research methodology, a summary of the articles of this dissertation, discussion of various aspects of the PhD research, and conclusion. Part II consists of two peer-reviewed journal articles, and one article that has recently been submitted for review.

Chapter 2 provides context for the research. It begins by introducing the field of learning analytics. Thereafter, data used within education and LA, emphasizing learning activity data and contextual data, are examined. Furthermore, educational data standards are explored, emphasizing xAPI and how context can be represented with this standard. Finally, LA architectures that utilize standards to integrate data and learning activity data storage are considered. Chapter 3 describes the methodology, starting with the overarching framework of design science research. Next, the research activities are outlined; first as they relate more broadly to the design science research framework, then the systematic literature review and case study that were applied within this framework are described. Chapter 4 gives an overview of the three studies that have been carried out in this research. Chapter 5 evaluates the research questions related to the contributions and discusses related research. Subsequently, methodological considerations and limitations of the research are presented. Chapter 6 concludes the dissertation by summarizing the research and pointing out future research directions.

2. Background

This chapter provides a backdrop for the research conducted as part of this dissertation. Initially, this chapter gives an introduction to the learning analytics field. Then the current state of educational data, emphasizing learning activity data and context data related to education, is reviewed. Next, standardized approaches to manage educational data are examined. Subsequently, further information related to LA architectures that utilize standards to integrate data are presented. Finally, research gaps related to LA scalability are summarily presented.

2.1 Learning Analytics

Ferguson (2012) details the history and drivers related to learning analytics. These include 1970s institutional research on student progression and persistence, the advent of the World Wide Web leading to new opportunities for collaborative learning and early educational platforms (e.g., LMS) in the late 90s, and the early 2000s with Web 2.0, broader LMS adoption, and new possibilities for dataset analysis.

In the introduction the most common definition of learning analytics, coined by George Siemens (Siemens, 2011), was given. This definition emerged from the first international conference on learning analytics and knowledge (LAK), which took place in Banff, Canada from February 27th to March 1st 2011. The definition was later adopted by the Society for Learning Analytics Research (SoLAR)[5], which was established in the summer of 2011 to manage the LAK conference, provide research directions for LA, and to promote the use of LA (Siemens & Baker, 2012).

Other definitions have also been given regarding learning analytics. For instance, George Siemens gave an early definition of learning analytics in 2010 when referring to it as: "The use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning" (as cited in Winne, 2017, p. 241). A somewhat simpler definition was given by Duval

[5] https://www.solaresearch.org/

(2012), who stated that learning analytics “is about collecting traces that learners leave behind and using those traces to improve learning” (Duval, 2012). Chatti et al. (2012, p. 322) defined learning analytics as a “research area that focuses on the development of methods for analyzing and detecting patterns within data collected from educational settings, and leverages those methods to support the learning experience”.

While there are variations in the different learning analytics definitions, it can be seen that the definitions all emphasize the use of educational data, which are analyzed by means of computational methods, where the results of the analysis methods are utilized to support or improve the learning experience.

2.1.1 LA process

The process of LA has been conceptualized in different ways (e.g., Chatti et al., 2012, Clow, 2012; Greller & Drachsler, 2012). The LA process model by Chatti et al. (2012), presented in Figure 2, comprises three main steps. As seen, data are first collected and pre-processed, subsequently there is some analytics and action, and finally post-processing may take place. These three steps may repeat iteratively, as lessons are learnt and refinements made to the process. The data are related to the educational domain, and data pre-processing involves steps such as data cleaning, transformation, and integration of data from multiple heterogeneous sources. Analytics are utilized to discover unknown patterns to support the learning experience, often supported by visualization, which may lead to taking action through activities such as recommendation, personalization, prediction, monitoring, and intervention. Finally, post-processing allows for making improvements to the analytics, for example through adding to or refining of the existing dataset, refinements in the selected variables/metrics, and selection of new analysis methods (Chatti et al., 2012).

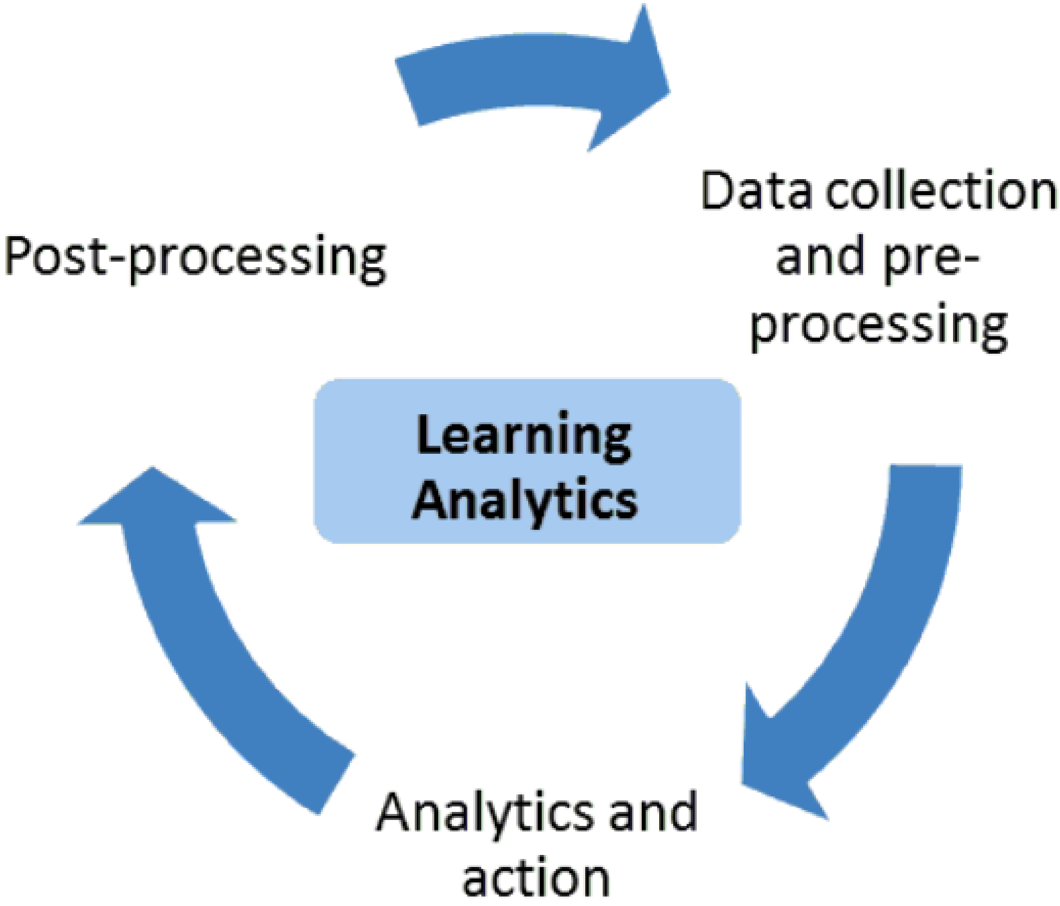

Figure 2 - Learning Analytics Process (Chatti et al., 2012)

While the process model by Chatti et al. (2012) highlights steps related to data and analytics, the LA life cycle by Khalil and Ebner (2015), shown in Figure 3, gives a more detailed illustration of the LA process, also emphasizing the human element. Here, it can be seen that the data collected are often big data about the learners, potentially from a multiplicity of sources, which may for instance derive from datasets containing interaction logs, academic information, traces from web activities, and further information pertaining to the students, for example personal data, and their learning context (Chatti et al., 2012; Khalil & Ebner, 2015). Subsequently, the data are analyzed according to techniques that may for instance relate to statistics, visualization, and social network analysis. Action may be taken aiming to optimize the learning environment, for example digital learning environments such as LMS, Massive open online course (MOOC) and simulation, according to objectives following from the interpretation of data and analysis. This illustration of feedback to the environment in the life cycle accommodates the concept of "closing the loop" suggested by Clow (2012), meaning that some intervention will result from the analysis. The intervention could for instance be related to a dashboard aimed at the individual students that shows their progression through a course and thus supports learner metacognition, or prediction indicating that a student is at risk of dropping out leading a student support employee in the administration at a higher educational institution to reach out to the student (Clow, 2012; Khalil & Ebner, 2015; Sclater et al., 2016).

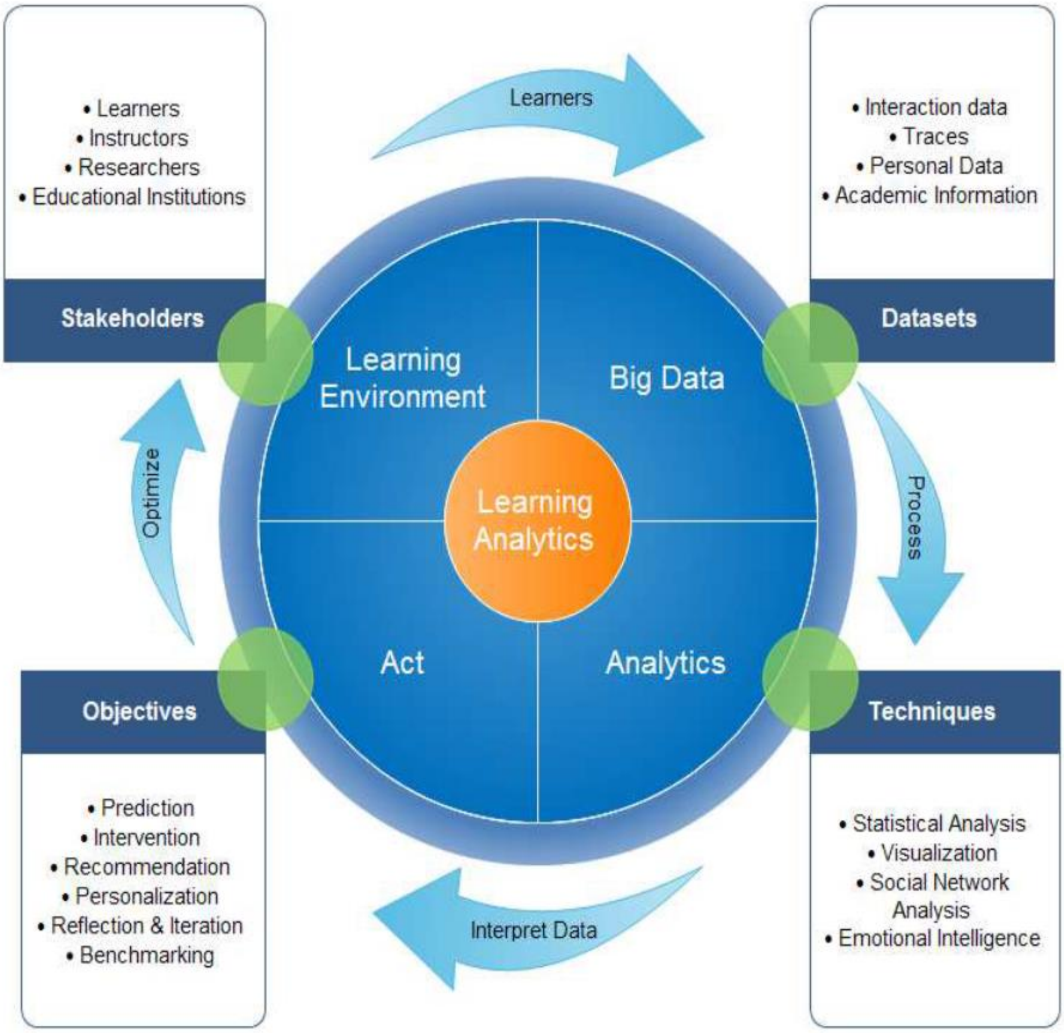

Figure 3 - Learning Analytics Life Cycle (Khalil & Ebner, 2015)

As shown in the life cycle by Khalil and Ebner (2015), there exists a number of possible stakeholders for learning analytics, including the learners themselves, the instructors that support the learners, researchers that are working within the learning analytics field, and the educational institutions (e.g., within K-12 or higher education) (Khalil & Ebner, 2015). As such, learning analytics may operate at both micro, meso and macro levels, where the micro level relates to the individual learner, the meso level relates to the institution, and the macro level concerns cross-institutional analytics, for example at the national or international level (Buckingham Shum, 2012). Furthermore, Greller and Drachsler (2012) illustrate how a division can be made for the stakeholders, between data subjects and data clients. Here, the data subjects are the individuals from whom the data are being generated and collected, typically the learner, while the data clients are stakeholders who gain information from the LA process, interpret and put in place interventions (e.g., teacher or administrative staff). In the case where the analytics results are fed back to the students they may simultaneously be data subjects and data clients.

In conjunction with their learning analytics life cycle, Khalil and Ebner (2015) also present a number of constraints that are relevant for the learning analytics process. More broadly, such constraints relate to privacy, ethics and data ownership, imposing limits for the application of learning analytics. Legislative frameworks, for example the General Data Protection Regulation (GDPR), inform institutional policy, giving regulations for conducting learning analytics in a more ethical, legal, and responsible manner (Khalil & Ebner, 2015).

2.2 Educational data

Data in the educational domain can be generated by a wide variety of sources. For instance, digital trace data may be generated by sources such as learning management systems, student admission systems, and social network systems (Hakimi et al., 2021). Digital tools and systems may facilitate data collection related to learning and education through questionnaires. Additionally, data from real-world (physical) learning contexts may be captured through devices containing sensors, including camera, microphone, eye-tracker, and electroencephalogram (EEG) headset (Ochoa & Worsley, 2016; Shankar et al., 2018; Giannakos et al. 2019). As Siemens and Long

(2011) point out, this context of data abundance can be described by the term “big data”. Big data may be defined as “datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze” (Manyika et al., 2011, as cited in Khalil & Ebner, 2015). Furthermore, big data is characterized by qualities such as their size (volume), real-time or close to real-time creation (velocity), and their availability in varying levels of structure (e.g., unstructured versus structured data; variety). Moreover, such data are typically relational, meaning data from different datasets may be connected (integrated) through common fields, and they are flexible for example with regard to accommodating new fields and increases in size (Kitchin & Lauriault, 2015).

Maintaining privacy is a crucial topic for LA and LA scalability. Privacy is one of the aspects of data security, where maintaining confidentiality, integrity and availability of the data are all of great importance (Kitchin, 2014, p. 175; Jain et al., 2016). De-identification of data adds to privacy, but when integrating data from a multiplicity of sources there is an added risk of persons (data subjects) being re-identified. Additionally, maintaining data privacy may prevent access to data that are of importance for LA research (Flanagan & Ogata, 2017). As such, different considerations related to LA and data privacy need to be taken into account, adhering to relevant laws.

2.2.1 Learning activity data

Data relevant for the educational domain, and used for LA, include data about learners as they are interacting with learning objects within (typically) digital learning environments (i.e., learning activity data), and data regarding learning content and services (Zouaq et al., 2017). While one may distinguish between learning activity data and learning content/service data, such data may also be represented collectively, or they may complement each other. For instance, the learning object that the learner is interacting with may be a type of learning content, in which case the learning content would also be a part of the learning activity data.

Activity data, including learning activity data, are related to the term “data exhaust”, indicating that these data are often generated as a byproduct of other processes, for example a user (or learner) interacting with a digital tool or product (Pomerantz, 2015). Since such data are often generated with a product-centric focus, for example stemming from log files for an (educational) tool or website, a criticism has been that they are not necessarily customized or the best suited for learning analytics (Kitto et al., 2020). Furthermore, activity data may be classified as either use data or use metadata, depending on the perspective, which are more generally related to the use of resources (e.g., learning objects) (Pomerantz, 2015; Baca, 2016, p. 10). Here, the metadata term is sometimes utilized since the data are stored separately from the process that triggered their generation. A common definition of metadata is simply “data about data” (Pomerantz, 2015, p. 19). As this definition makes clear, metadata are themselves data.

2.2.2 Context

In addition to information about the learner interacting with a learning object, it is also important to store contextual information related to the learning activity data, for instance to account for learning taking place in both formal and informal settings (Chatti et al., 2014; Bakharia et al., 2016). As such, learning activity data need to be enriched with further contextual data, and modeling of contextual information is itself a challenge.

As seen in the Introduction chapter, a common definition of context pertains to the situation of an entity (Dey, 2001, p. 5). The meaning of context may, however, vary according to perspective. A social and user-centered perspective on context is that it is a characteristic of an interaction, thus the context includes the factors that have an impact on a learner involved in a learning activity. To operationalize context in a more technical sense, researchers have defined context according to context dimensions, for example the device and environment context. Such context dimensions group a set of related context properties (Verbert et al., 2012). Utilizing context information for LA may further contribute to areas such as adaptivity, i.e., adapting learning to the individual learner, personalization, and recommendation (Chatti et al., 2014; De Meester et al., 2018).

2.3 Educational data standards

A number of standardization initiatives have been undertaken related to education, in areas such as representation of learning activity data, context models, learning object metadata, and learning design (Jovanović et al., 2007; Verbert et al., 2012; Lukarov et al., 2014; Shankar et al., 2022). Utilizing standards, for example to represent activity data and context, can help with regard to aspects such as data integration and interoperability (Cooper, 2014a), as mentioned in the Introduction chapter.

Interoperability can be considered from the semantic, technical, legal and organizational level (European Commission, 2017). The *semantic* level relates to the preservation of meaning and data formats, thus enabling proper understanding of the data among different parties. The *technical* level includes services for the exchange and integration of data. The *legal level* is about ensuring collaboration among different parties, even if they have different legal frameworks, strategies and policies. User privacy protection is important in terms of this interoperability level. Finally, *organizational* interoperability includes aligning business processes for achieving shared organizational goals and responding to needs of the users (European Commission, 2017).

In the domain of education, standards typically consist of multiple parts. A model often being utilized for EdTech standards comprises: 1) an abstract *data model* containing information related to the structure and organization of data; 2) *bindings* that specify how the data model is expressed formally (through machine-readable formats such as JavaScript Object Notation [JSON] or Extensible Markup Language [XML]); and, 3) *Application Programming Interface (API)* that enables contact points for machine-to-machine communication (Friesen, 2005). Furthermore, standards may pertain to *data values*, for example by specifying the type of values that are permitted or rules for *content* that can be used to populate the data structures (Baca, 2016, p. 3).

Despite their potential for data integration and interoperability, the creation of different standards and vocabularies to represent similar types of data may lead to fragmented sub-communities without cross-standard interoperability (Zouaq et al., 2017).

2.3.1 Standards and specifications

It may be difficult to differentiate between the terms specification and standard. One relevant aspect in differentiating between specifications and standards is that accredited standards bodies, such as International Organization for Standardization (ISO) and Institute of Electrical and Electronics Engineers (IEEE), may create official standards, while consortia including 1EdTech (formerly IMS) may create technical specifications (Duval, 2004; Sclater, 2017). Development of specifications may follow a less rigorous process than standard development, for instance accepting inputs from multiple stakeholders and allowing for development to be a work in progress (Sclater, 2017). Within consortia, there is emphasis on the requirements and needs from member organizations. This is in contrast to the process of accredited standards bodies, which have a broader focus in terms of the needs of all stakeholders (in contrast to only member organizations) and that aim for a fair and open process (Duval, 2004). Once a specification is more clearly defined, stable, and unambiguous in terms of implementation, having been through experimentation and validation, it may evolve into an official standard after being adopted by a standards body (Duval, 2004; Sclater, 2017). Another aspect is that even though a technical specification is not released by a standards body, and it may still be at a fluid state of development, it can still become a *de facto* standard, meaning it gains real-world adoption. Additionally, there is the notion of a *de jure* standard, which implies the standard has been developed through an open process, typically in an accredited standards body (Duval, 2004). As noted by Cooper (2014b) the use of standard versus specification terminology is less uniform outside of the formal standardization process. When the term standard is used in this dissertation, it may refer both to official standards and de-facto standards.

2.3.2 Linked data technologies

Linked data technologies (also referred to as semantic web technologies), which enable generic data representations within and across domains, may be used as a means to provide a standardized representation of educational data, thus enabling integration of heterogeneous sources. Such technologies may be, and have been, adopted within individual EdTech standards communities, for example to provide clearer semantics of terms (i.e., vocabulary concepts and relationships), and furthermore they allow bridging of terms across standards communities (Zouaq et al., 2017).

A number of World Wide Web Consortium (W3C) standards may be utilized for data representation (Zouaq et al., 2017). A central component for linked data is the Resource Description Framework (RDF), which can express and connect data about resources (entities), and which is "the basic framework that the rest of the Semantic Web is based on" (Allemang et al., 2020, p. 33). RDF enables the use of data types that are largely taken from the XML Schema Definition (XSD). To add meaning to linked data it is possible to use modeling languages to represent vocabularies and/or ontologies, which provide a basis for describing entities and their relationships, through standards that build on and are expressed using RDF, such as RDF schema (RDFS), Web Ontology Language (OWL) and Simple Knowledge Organization System (SKOS) (Bizer et al., 2011; Allemang et al., 2020). In addition to modeling languages for definition of vocabularies and ontologies, there also exists linked data standards that allow for modeling of expectation; for instance, the Shapes Constraint Language (SHACL) enables the modeling of expectations for a certain type of data, both at the level of what is to be expected from a resource (e.g., what properties[6] should be included) and at the level of a property (e.g. what should be the data type of a value). Following, this language can be used, among other things, for data validation (Allemang et al., 2020).

A number of different serializations are available for expressing RDF. A more recent serialization is the JSON for Linking Data (JSON-LD)[7] serialization. JSON-LD allows for the RDF data to be expressed using JSON. A central motivation behind JSON-LD is that developers who are familiar with JSON, but not with linked data standards, may nevertheless be able to build linked data compatible applications. Using the JSON-LD @context keyword it is possible to define or include rules for how the JSON-LD document should be interpreted, for instance mapping of terms to their corresponding Internationalized Resource Identifier (IRI)s (Allemang et al., 2020).

[6] In linked data terminology a property is often referred to as a predicate.
[7] https://www.w3.org/TR/json-ld11/

Despite their potential related to data integration, the uptake of linked data standards and concepts in the LA community has been quite slow (Zouaq et al., 2017).

2.3.3 Context data models

Different data models have been developed in the educational domain to represent context, including Learning Object Context Ontologies (LOCO), Contextualized Attention Metadata (CAM), the Learning Context Data Model (LCDM), and the rich context model (RCM) (Jovanović et al., 2007; Schmitz et al., 2011; Thüs et al., 2012; Lincke, 2020). LOCO, a framework based on ontologies, was developed to allow for representation of contextual information regarding the use of digital learning objects (Jovanović et al., 2007). The CAM framework allows for collection of data related to users' selection behavior in digital settings, for example regarding data objects given attention by the users (Schmitz et al., 2011). The LCDM, which takes learners and events into consideration, enables the representation of both intrinsic and extrinsic context. Here, intrinsic context relates to the inside of the learner (e.g., their interests), while external context relates to the learner's environment (Thüs et al., 2012; Thüs et al., 2015). In the RCM, the context of the user is modeled according to context dimensions that individually relate to an application domain, for example mobile learning and LA (Lincke, 2020). The context data models vary in aspects such as their unit of focus (e.g., learner or learning resource), flexibility of registering data, how the context is categorized (flat vs. hierarchical structure), and representation format. More detail on these models is provided in Paper 2.

2.3.4 Learning activity data standards

This section provides information related to the learning activity data standards xAPI and Caliper. The emphasis will be on xAPI, including statements and profiles, since the research detailed in Paper 2 and Paper 3 has focused on extending xAPI for enhanced context descriptions, as part of the case study that examined xAPI from the perspective of the AVT project. Additionally, some of the projects that precede and complement xAPI are reviewed to gain a broader understanding of the backdrop of xAPI and its place as part of a broader architectural vision.

xAPI and Caliper are both learning activity data standards that enable description, integration and exchange of learning activity data. Furthermore, they both enable the definition of vocabularies through profiles that may thereby help to clarify the semantics of the activity data. Profiles may also provide a number of rules, for example in terms of the allowed activity data syntax related to a specific use-case or domain. The profiles for xAPI and Caliper are both represented with the JSON-LD format, and are thus specified in the machine-readable JSON format, which has allowed for adoption of linked data concepts (Griffiths & Hoel, 2016; Zouaq et al., 2017).

To format activity data according to a learning activity data standard, such as xAPI and Caliper, certain tools, systems, and platforms related to education provide data export capabilities in the given format, for example through plugins (Betts & Smith, 2019). As pointed out by Kitto et al. (2020, p. 13), “large vendors usually support at least one specification and sometimes both”. Alternatively, data in one format may go through transformation according to certain rules to end up in the specified format (Sclater, 2017). As such, work has been done within the International Organization for Standardization (2020) with regard to the technical specification “Learning analytics interoperability — Part 3: Guidelines for data interoperability”, to provide guidelines for syntactic and semantic mappings between what they refer to as LA data API specifications, using xAPI and Caliper as reference specifications. Upon learning activity data being generated they may be stored in a Learning Record Store (LRS)[8].

While there are a number of similarities among xAPI and Caliper, there also exist differences (1EdTech, n.d.-b; Griffiths & Hoel, 2016). The focus of xAPI is on tracking any type of learning experience, whether in the physical or digital realm, while Caliper analytics was developed to “support advanced learning measurements and analytics by providing recommendations on how to capture and share events data from learning systems which are based on existing and upcoming IMS [1EdTech] specifications” (Muslim, 2018, p. 55). xAPI statements, where a statement is the data structure that is used to represent xAPI activity data, are represented using JSON, while Caliper activity

[8] Storage of activity data in an LRS is not part of the Caliper Analytics standard, but the LRS has become a de facto storage solution also for Caliper activity data (Miller & Duan, 2018).

data are specified with JSON-LD. Profiles are a part of the core functionality of Caliper, while they are currently extensions in xAPI. For xAPI, the intention is that individual communities of practice will create profiles as needed for their specific purpose. xAPI, whose development is facilitated by Advanced Distributed Learning (ADL), is open source, meaning anyone can suggest standard changes and new xAPI profiles can be defined by any community of practice. Caliper analytics is developed by the 1EdTech consortium, and while the Caliper codebase is released as open source, member organizations are responsible for contributing to the standards development and for approving their publication. As such, it is the use-cases of the IMS membership organizations that are the main focus of the specification (Griffiths et al., 2016; Griffiths & Hoel, 2016; Kitto et al., 2020). Consequently, xAPI generally offers more flexibility in its use than Caliper analytics. A more detailed comparison of the two specifications can be found in Griffiths and Hoel (2016).

2.3.4.1 xAPI

Related to xAPI, both the xAPI specification (Advanced Distributed Learning, 2021b) and the xAPI profile specification (Advanced Distributed Learning, 2022a) exist. The xAPI specification, currently in version 1.0.3, provides the general structure in which *all* xAPI users can describe learning activity data in the form of an xAPI statement. Complementing the xAPI specification that focuses on the xAPI statement structure, the xAPI profile specification mainly concerns itself with the content of the statements, allowing communities of practice to define their own vocabulary concepts and rules.

Figure 4 shows the basic building blocks of an xAPI statement (Vidal et al., 2015). As mentioned in the Introduction chapter, xAPI has the most basic structures of actor, verb, and object, which are all required, and furthermore allows for the registration of contextual data. Contextual data may be placed within the xAPI context and results structures, which exist on the same level as the actor, verb, and object structures. While the results structure regards a measured outcome connected to the xAPI statement where it exists and is thus primarily related to assessment, the context structure concerns the learner and their interaction with a learning object (Advanced Distributed Learning, 2021c).

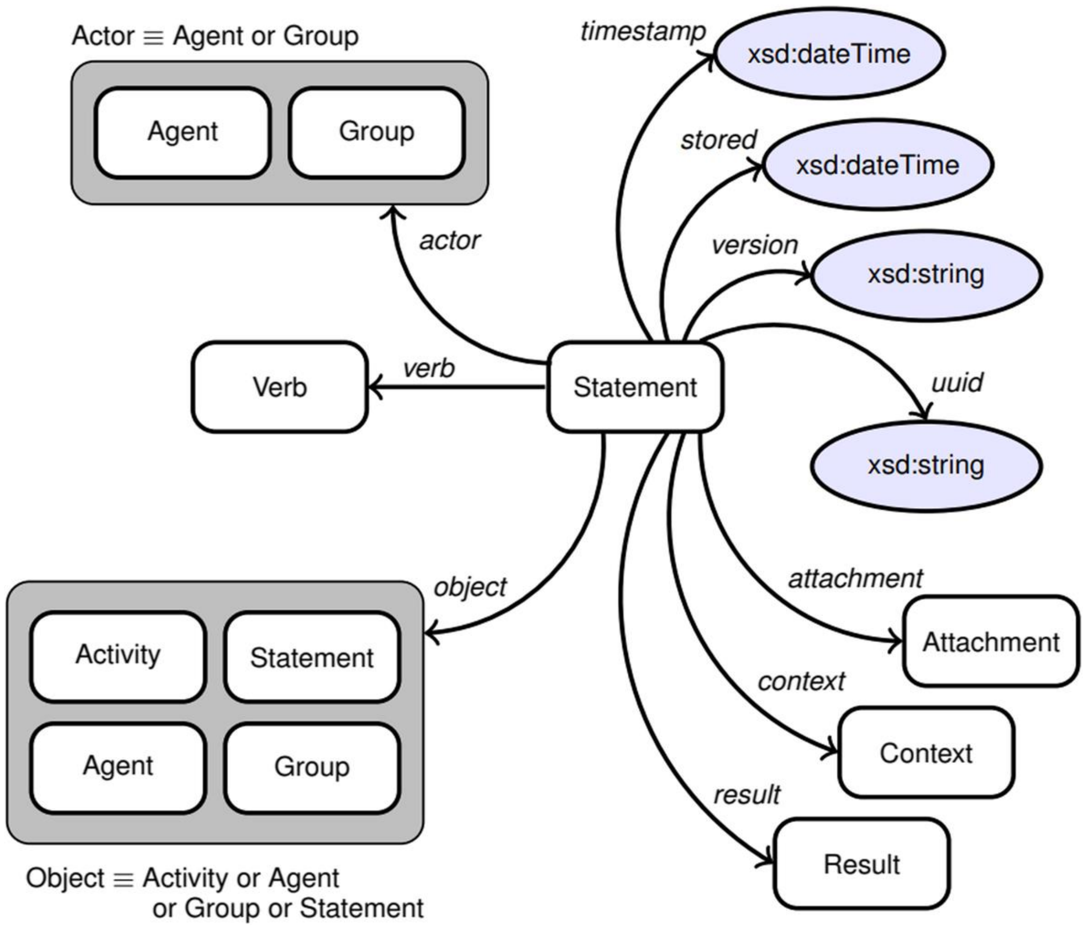

Figure 4 - xAPI statement (Vidal et al., 2015)

Typically, the object of an xAPI statement, i.e., the object that a learner (the agent) has interacted with, is an xAPI Activity. The verb and Activity may collectively represent "a unit of instruction, experience, or performance" (Advanced Distributed Learning, 2017a). Furthermore, the Activity can represent both tangible and virtual objects. Further detail on the xAPI statement is given in Paper 2.

2.3.4.2 xAPI context

Within the xAPI context structure, a number of properties may be defined (see Table 1). Seven of the properties take a single object or value, while the remaining 2 properties, i.e., contextActivities and extensions, which are of type Object, allow for registration of more than one entity or value.

Table 1 - Context structure properties (Advanced Distributed Learning, 2021c)

Property	Type	Description	Required
registration	UUID	The registration that the Statement is associated with.	Optional
instructor	Agent (MAY be a Group)	Instructor that the Statement relates to, if not included as the Actor of the Statement.	Optional
team	Group	Team that this Statement relates to, if not included as the Actor of the Statement.	Optional
contextActivities	contextActivities Object	A map of the types of learning activity context that this Statement is related to. Valid context types are: "parent", "grouping", "category" and "other".	Optional
revision	String	Revision of the learning activity associated with this Statement. Format is free.	Optional
platform	String	Platform used in the experience of this learning activity.	Optional
language	String (as defined in	Code representing the language in	Optional

	RFC 5646)	which the experience being recorded in this Statement (mainly) occurred in, if applicable and known.	
statement	Statement Reference	Another Statement to be considered as context for this Statement.	Optional
extensions	Object	A map of any other domain-specific context relevant to this Statement. For example, in a flight simulator altitude, airspeed, wind, attitude, GPS coordinates might all be relevant.	Optional

ContextActivities allow for the registration of xAPI Activities as part of the xAPI context. The Activities may be grouped within the contextActivity types grouping, category, other, and parent. As such, there are different ways to register the same type of context data using contextActivities. (Context) extensions, on the other hand, allow for registration of any context data relevant for the statement that are not covered by the other context structure properties and are not directly related to a result. As such, the extensions are flexible, allowing for definitions of arbitrary key-value pairs, where the values must be within the JSON data types including string, array, and object. Furthermore, other context properties, such as platform, do not specify format requirements for the string value. As such, xAPI is quite flexible in terms of context registration. The flexibility of xAPI context, i.e., the variety in ways contextual information may be represented, has been found to be a challenge in terms of combining the data (Bakharia et al., 2016; Kitto et al., 2020). Additionally, xAPI utilizes a flat structure for context data (rather than hierarchical - e.g., 2 level categorization). Further details on the xAPI context structure, and a comparison between xAPI context and the data models listed in the “Context data models” section, can be found in Paper 2.

2.3.4.3 Profiles, validation and developments

The xAPI profile specification includes a number of constructs that enable the definition of vocabulary concepts and rules (Advanced Distributed Learning, 2022b), and furthermore the specification utilizes several more generic specifications. When defining xAPI extensions in a profile, for use in xAPI statements that conform to the profile, it is possible to include a JSON schema[9] using the extensions "schema" or "inlineSchema" property, allowing for a description of the extensions and for definition of constraints. JSON Schema supports constraining JSON structures according to the six basic data types available in JSON, i.e., string, number, boolean, object, array and null. For each of these types further constraints are available, for instance the type *string* can be used in conjunction with a regular expression pattern or a defined format (e.g., date, date-time) to specify the allowed format for the string, and the type *object* can be used in conjunction with the *properties* and *required* keywords to specify characteristics of properties within an object and which of these properties are required.

Other profile constructs that may be used to constrain data in xAPI statements that adhere to a profile include the statement template and pattern constructs. Statement templates may be utilized to specify which data must be present in an xAPI statement that describes a given event. Statement template rules are partly specified using the JSONPath[10] specification. Patterns can define the ordering of a collection of statements that are defined by statement templates (Advanced Distributed Learning, 2022b).

The xAPI profile specification reuses a number of vocabulary constructs from semantic web vocabularies, such as Simple Knowledge Organization System (SKOS), thus

[9] https://json-schema.org/
[10] https://goessner.net/articles/JsonPath/index.html

adding structure and semantics to the community-defined profiles. A number of mapping properties are also reused, through the use of SKOS, to allow linking of concepts across vocabularies. Additionally, it has been proposed that the xAPI specification could incorporate linked data concepts by utilizing the JSON-LD format to allow exposing the learning activity data as linked data[11] (De Nies et al., 2015). This has, however, not yet been implemented as part of the xAPI specification.

Tools that "validate the conformance of data to [...] specifications [such as xAPI] are key components to assure the interoperability among applications" (Rabelo et al., 2017, p. 1). In the case of xAPI, validation may be provided on the levels of statement, profile, and statement according to the profile.

The xAPI specification is set to be released in version 2.0 as an IEEE Learning Technology Standards Committee (LTSC) standard, i.e., the xAPI base standard, through efforts by the IEEE xAPI Working Group (Advanced Distributed Learning, 2020; IEEE, 2022). The IEEE xAPI Working Group will also standardize xAPI profiles, through several companion standards that will complement the xAPI base standard. This IEEE effort for profiles, however, is at an earlier stage than for the xAPI base standard (IEEE, 2022).

2.3.4.4 xAPI - Related projects, specifications and standards

xAPI is part of the Total Learning Architecture (TLA), an ADL project "to develop a set of technical specifications, standards, and policy guidance that define a uniform approach for integrating current and emerging learning technologies into a *future learning ecosystem*" (Gordon et al., 2020, p. 2). To integrate data related to learning, the project aims to scale across temporal, institutional and technical boundaries, based on four pillars for integrated data. These pillars relate to: 1) Learning activity tracking; 2) experience index of learning objects and their content; 3) competency framework; and, 4) learner profiles (Gordon et al., 2020).

[11] https://github.com/adlnet/xapi-ontology

xAPI is related to the first TLA pillar for integrated data, i.e., learning activity tracking. In terms of description and storage of learner records, xAPI is an improvement from the earlier e-learning specification called Sharable Content Object Reference Model (SCORM), which was developed by ADL, first being released in 2001, and which could only track a limited number of learning events (related to completion and scores) that occurred as a learner interacted with training content in a web-browser based learning management system (LMS; Griffiths et al., 2016). SCORM, however, did not only address tracking of learning activity data; it also addressed, for example, launch and packaging of re-usable e-learning content units, and enabled run-time communication between the LMS and content units (Bailey, 2005). Thus, xAPI is only an improvement to, and replacement of, the tracking aspect of SCORM.

Another specification that addresses learning activity tracking, focusing on learning activities in the LMS, is the cmi5 specification[12]. The cmi5 specification is represented through an xAPI profile, thus adhering to the xAPI standard, but cmi5 also includes rules for aspects such as interoperable content launch, content packaging, and communication between LMS and content. As such, cmi5 both utilizes and complements xAPI as a future replacement for SCORM. A final specification that is also represented using an xAPI profile is the TLA Master Object Model (MOM)[13]. It serves to normalize the xAPI activity data (i.e., xAPI statements) that are used as input to TLA services (e.g., related to competency management services, which collect evidence in the form of xAPI statements to make competency assertions regarding individuals and teams), by providing a controlled vocabulary that the xAPI statements must conform to (Gordon et al., 2020).

While xAPI also records data regarding learning objects, the unit of focus for xAPI is the learner who engages in a learning activity. There are complementing standards that focus on learning objects, and thus allow for more structured registration of this type of information. These standards include Learning Resource Metadata Initiative (LRMI) and IEEE Learning Object Metadata (LOM) (Gordon et al., 2020). The Standard for

[12] https://github.com/AICC/CMI-5_Spec_Current
[13] https://github.com/adlnet/MasterObjectModel

Learning Metadata[14], which is currently a draft standard, aims to accommodate the changing landscape for educational technologies. This draft standard builds on LOM, but also harmonizes with other educational data standards, including LRMI (Smith & Milham, 2021). Focus areas include the provision of a rich vocabulary and adding granularity for learning object descriptions (Advanced Distributed Learning, 2021a). Related to the TLA developed by ADL, metadata about learning objects is stored in the experience index, where the metadata descriptions are in accordance with the Standard for Learning Metadata draft standard (Smith & Milham, 2021). This makes it possible to enhance xAPI learning object data with additional metadata from the experience index.

2.4 Learning analytics architectures and storage

A number of learning analytics architectures have been developed, for example from Apereo, JISC, and SURFnet (Muslim, 2018). At the technical level, these architectures more generally have capabilities for collection, analysis and reporting of educational data, related to specific objectives, such as personalization, intervention, and monitoring (Sclater, 2017; Muslim, 2018). The architectures facilitate integration of data originating from different sources, for instance through plugins, connectors, or data transformations (Berg, Scheffel et al., 2016; Apereo, 2020; JISC, 2020). The LA architectures comprise discrete components that may be reusable. For instance, the OpenLRS, which is a storage component of the Apereo architecture, has been utilized in a learning analytics platform that has been deployed at a Japanese higher education institution (Flanagan & Ogata, 2018).

Both the Apereo, JISC, and SURFnet architectures utilize xAPI for storage of learning activity data, meaning the data can be combined in an LRS. Additionally, Apereo has also introduced support for Caliper activity data, through storage in an LRS with added capabilities for handling certain types of educational data that are not learning activity data, i.e., the learning record warehouse (Sclater, 2017; Muslim, 2018). An (xAPI) LRS

[14] https://standards.ieee.org/ieee/2881/10248/

builds on existing storage technology, such as a noSQL database or a relational database technology, but additionally implements the LRS RESTful interfaces according to the xAPI specification (Advanced Distributed Learning, 2017b). Advanced Distributed Learning keeps a list of LRSs that are conformant to the standard[15]. Furthermore, it is possible to store the learning activity data in an LRS utilizing an RDF store as its storage backend, which is tuned for RDF data storage and retrieval (Anseeuw et al., 2016; Allemang et al., 2020, p. 70).

2.5 Research gaps

As this and the preceding chapter have shown, there are gaps in research related to LA scalability that need to be filled.

Data integration, which is closely linked to interoperability, is an important aspect that contributes to LA scalability. However, data that originate from various sources are often represented using different formats, with varying levels of structure, and stored using diverse storage technologies. This leads to a situation where data exist in silos, hindering integration and analysis across sources. As such, data integration is a significant challenge in the field of LA. While utilizing standards such as xAPI may contribute to data integration in the educational domain, previous research has identified that there are issues with xAPI and combining data. Among the issues is that the flexibility of xAPI context means that contextual information may be described in a multitude of ways, which may hinder consistent data descriptions across sources (i.e., expressibility).

The research presented in this dissertation aimed to explore and address challenges for LA data integration, interoperability, and consequently scalability. More specifically, it focused on identifying gaps in xAPI expressibility and providing solutions to these gaps, focusing on xAPI context.

[15] https://adopters.adlnet.gov/products/all/0

3. Methodology

This chapter describes the methods used to conduct the research and explains how they are applied in the research. First, the design science research (DSR) framework will be introduced. Next, the separate activities conducted as part of this research are explained, first on a broader level related to the design science research framework, and subsequently as they relate more specifically to the systematic literature review and case study research methods that were utilized within this framework.

3.1 Design science research

Design science research (DSR) was used as the overarching research framework for this research. DSR, a problem-solving paradigm that has origins in engineering and the sciences, is concerned with the construction of novel and innovative artifacts aiming to solve real and important problems. This relates to the goal of DSR, which is utility through problem solving (Hevner et al., 2004). DSR was chosen as a research framework as it fits the aims of the research, which involved stages of 1) gaining an understanding of LA scalability challenges through examining the current state of data integration and interoperability, 2) identifying gaps and needs in a current learning activity data standard (i.e., xAPI) to build requirements (recommendations) for its improvement that were later implemented as part of a technical solution, and 3) validating the technical solution through rigorous evaluation. The research framework accommodates the different research strategies and activities chosen to pursue the aims of this research. By anchoring the research in the DSR framework the problem domain could be studied and followed up with the building of artifacts, which were later subject to evaluation, thus addressing identified gaps and needs.

Figure 5 displays a conceptual framework for carrying out research related to design science (Hevner et al., 2004). As shown, research is conducted where artifacts are developed/built and evaluated. The artifacts developed stem from a need in the environment (based on people, organizations and technology), and must also be (re-)applied to the environment, ensuring relevance. Rigor assures that the DSR, related to building and evaluation of the artifacts, is grounded on strong foundations from the knowledge base (Hevner et al., 2004). These foundations are not only related to scientific theory and methods, but also to experience, expertise, artifacts and processes that exist in the environment, where the outputs of design science research may also add contributions back to the knowledge base (Hevner, 2007).

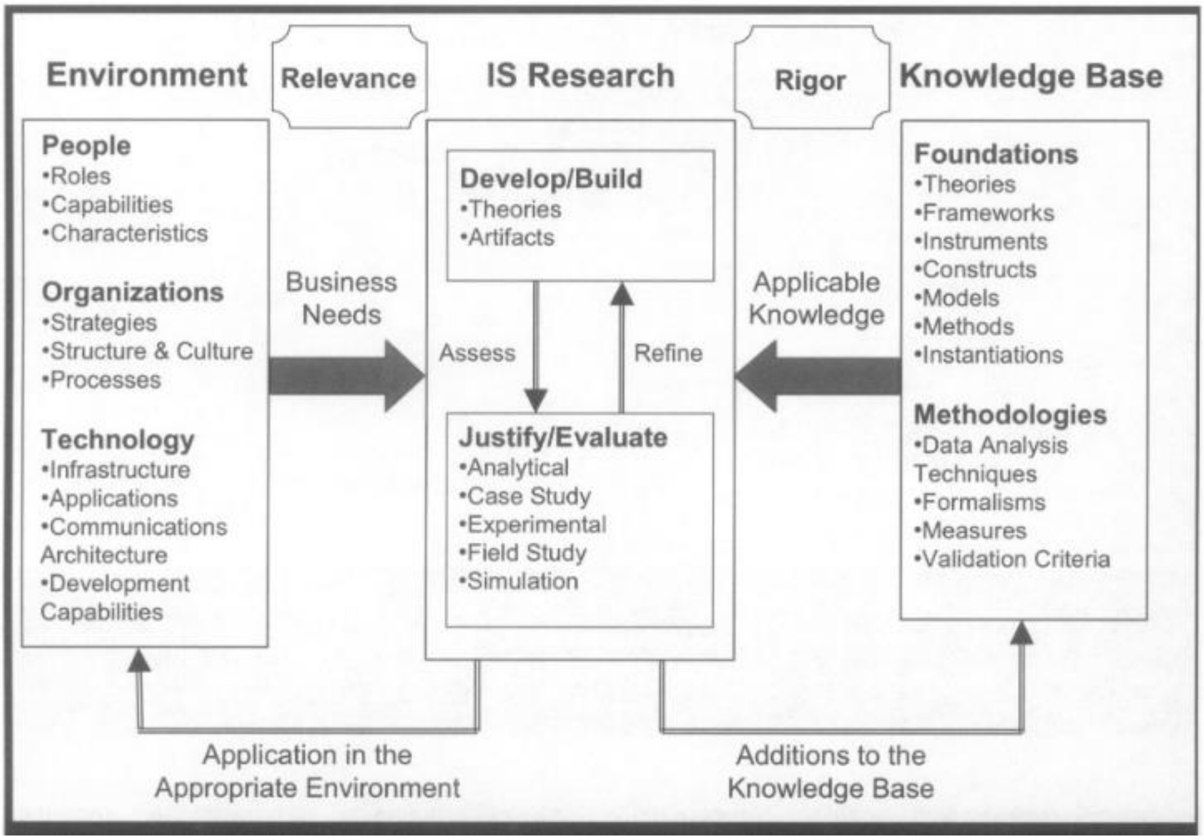

Figure 5 - Information Systems Research Framework (Hevner et al., 2004)

A central concept in DSR is that of the artifact, i.e., an artificial (man-made) object that may constitute a set of prescriptions that can be applied by researchers and practitioners to understand and handle the inherent challenges encountered when designing and developing information systems in organizational environments, and which may be either abstract or material (instantiated) (Vaishnavi & Kuechler, n.d.). Building on artifact categorizations from the foundational DSR paper by March and Smith (1995), Hevner et al. (2004, p. 77) broadly define artifacts as "*constructs* (vocabulary and models), *models* (abstractions and representations), *methods* (algorithms and practices), and *instantiations* (implemented and prototype systems)". In more recent years, other typologies have been provided for accepted DSR artifacts (e.g., Vaishnavi & Kuechler, n.d.; Offermann et al., 2010). Based on an extensive literature review of DSR, Offermann et al. (2010) include additional artifact types such as *requirements* (statements about a system regarding the types of properties it should have, and [optionally] the reason why), and *pattern* (generic reusable system design elements, described according to its benefits and applicable context). The artifacts are built in an iterative manner, based on experiences and knowledge of the problem area and its solution, which is gained through iterations of building followed by evaluation.

Evaluation ensures artifacts are assessed with regard to the utility they provide in problem solving. For evaluation, metrics must be defined that operationalize what constitutes utility related to the developed artifact, and typically data appropriate for the type of evaluation are collected and analyzed. A variety of evaluation methods that relate to DSR are available in the knowledge base. The chosen method, however, must be appropriate for the developed artifact and the chosen metrics (Hevner et al., 2004).

As Pries-Heje et al. (2008) point out, evaluation in DSR may be viewed from two different viewpoints. The first viewpoint is the *ex ante* perspective, where artifacts are evaluated before implementation/actual construction, while the second viewpoint is the *ex post* perspective, where artifacts are evaluated after design/construction. Ex post evaluations have traditionally been the commonly assumed evaluations in DSR processes, implying a clear distinction between the construction and evaluation activities (Sonnenberg & vom Brocke, 2012). Sonnenberg and vom Brocke (2012) details how throughout the construction process, the artifact will *emerge* through interactions with the environment (organizational context), and how knowledge building and reflection will also alter the artifact design. In contrast to ex post evaluations, the ex ante evaluation accommodates this emergent process with construction and evaluation activities that may happen in parallel, meaning the progress achieved may be reflected without delay, and that artifacts may be modified early in the design process. The ex ante and ex post distinction has traditionally been associated with information systems research, assuming that it is a system that is being constructed/implemented, whereas a DSR artifact may also, for example, relate to requirements or design. Consequently, the ex ante and ex post distinction may slide according to the definition of the artifact (Pries-Heje et al., 2008; Sonnenberg & vom Brocke, 2011).

Evaluations in DSR may be classified as either artificial or naturalistic (Pries-Heje et al., 2008). The naturalistic evaluation takes place in a real-life environment, i.e., the organization, acknowledging the complexities that may arise when humans interact with technologies in authentic environments, and will involve “real users using real systems to solve real problems (i.e., to accomplish real tasks in real settings)” (Pries-Heje et al., 2008, p. 4). Venable (2006) highlights the importance of having artifact evaluations in a naturalistic environment. In contrast to the naturalistic evaluation, the artificial evaluation proceeds in a contrived way with abstraction leading to less involvement with the natural setting (Pries-Heje et al., 2008). Regarding ex ante evaluations, the lack of a developed artifact that is applied to a real problem means that these evaluations are necessarily artificial. Ex post evaluations, on the other hand, may be either naturalistic or artificial, depending on what is the aim of the evaluation and what evaluation methods are used (Sonnenberg & vom Brocke, 2011).

3.1.1 Philosophical paradigm

Regarding the philosophical paradigm used for design science research, where the term paradigm may be used to characterize the researcher's “worldview” (Kivunja & Kuyini, 2017), multiple discourses have emerged, including interpretivism, positivism, and pragmatism (van der Merwe et al., 2020). The worldview, as viewpoint, thinking or shared beliefs, affects decisions related to the research, such as which methods to apply and how data are understood or interpreted by the researcher (Kivunja & Kuyini, 2017). Within fields related to computing, interpretive research is concerned with “understanding the social context of an information system: the social processes by which it is developed and construed by people and through which it influences, and is influenced by, its social setting” (Oates, 2006, p. 292). Interpretivist research tries to understand phenomena by exploring how people perceive the phenomena in their own (natural) setting, acknowledging that perceptions may vary among people and across groups. As such, interpretivist research understands reality or knowledge to be subjectively constructed by the minds of individuals and groups, and therefore there is no single and objective truth. Meaning is constructed socially and dynamically, as whatever reality essentially is, it is only accessible and transmittable through social constructs such as language and shared understanding, which will differ between groups and is subject to change over time. Qualitative data are typically collected and analyzed within interpretivism, such as the words and metaphors employed by the users of an information system. Positivism is often associated with the natural sciences, and underlies the scientific method. It assumes the world is governed by laws, patterns, or regularities, which research aims to identify, through methods such as experiments and surveys. In this view there is an assumption that the world can be studied objectively, because the laws and patterns exist independent of humans and their personal experience. As such, there is a criticism that positivist methods may not be suited for studies of phenomena taking place in a social context, for instance related to individuals' thoughts and actions. Collected data are typically subjected to quantitative data analysis, with mathematical modeling and statistical analysis (Oates, 2006). As a reaction to interpretivism and positivism, the pragmatic paradigm “arose among philosophers who argued that it was not possible to access the ‘truth’ about the real world solely by virtue of a single scientific method as advocated by the Positivist paradigm, nor was it possible to determine social reality as constructed under the Interpretivist paradigm” (Kivunja & Kuyini, 2017, p. 35). Pragmatism, emphasizing research approaches that work in practice, therefore accommodates the use of the research methods that are deemed most appropriate based on the phenomenon being studied, whether they are typically associated with interpretivism or positivism (or another paradigm, for example the critical paradigm), and whether they collect qualitative or quantitative data. Thus, pragmatism allows multiple research methods to be used in conjunction, which may collectively elucidate human behavior and belief (Kivunja & Kuyini, 2017).

Through its emphasis on utility, DSR is concerned with “creating what is effective” (Hevner et al., 2004), meaning real-world problems should be solved through the DSR process. Hevner (2007) argues that DSR is thus aligned with the philosophical paradigm of pragmatism, which "considers practical consequences or real effects to be vital components of both meaning and truth" (Hevner, 2007, p. 91). According to Hevner (2007) it is, however, not only the relevance of the problem, but also the rigor of the process and the research contributions that constitute well-executed design science research. As mentioned earlier in this section, the pragmatic philosophy allows for different types of evaluation methods to be used within the research. This diversity of methods is reflected for design science research, following the many evaluation methods suggested and applied within this framework, for example case study or experiment (see Figure 5) (Hevner et al., 2004; Sonnenberg & vom Brocke, 2012).

3.1.2 DSR process

DSR process models may be used to describe and visualize the DSR process at a more general level, and a number of such models have been created (e.g., Vaishnavi & Kuechler, n.d.). The models have many similarities, but may emphasize different aspects of DSR. Vaishnavi and Kuechler (n.d.) developed their often cited process model originally in 2004, although the work has been updated in recent years. The model, which is shown in Figure 6, emphasizes the creation of new knowledge through following DSR, and includes the following process steps: Awareness of the problem, suggestion, development, evaluation, and conclusion (Vaishnavi & Kuechler, n.d.).

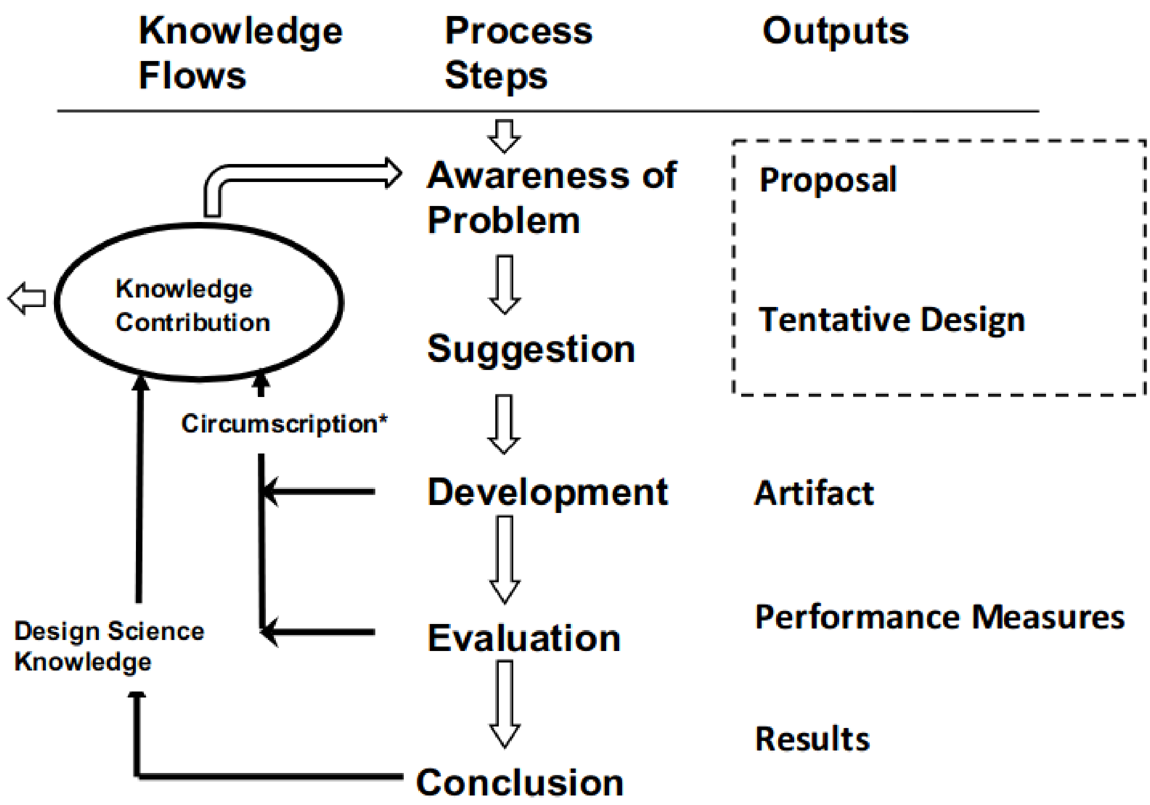

Figure 6 - Design Science Research Process Model (DSR Cycle)
(Vaishnavi & Kuechler, n.d.)

Awareness of the research problem may stem from different sources, including organizations, problems described in the research literature, or reading research literature in allied fields that present opportunities to apply recent findings to the field of the researcher. The problem awareness phase outputs a proposal, which may be either formal or informally stated, regarding a new research venture to pursue. Next, closely tied to the proposal, a suggestion is made for new functionality, which will typically result in a tentative design, but can also result in a proposed approach to develop the solution that is offering new functionality if the solution is not currently clear (Vaishnavi & Kuechler, n.d.; van der Merwe et al., 2017). During the development phase the tentative design or proposed approach will be further refined and constructed/implemented, resulting in an artifact. The evaluation phase includes evaluating the artifact according to applicable criteria. Typically the criteria will be related to the utility of the artifact, with careful collection of results and tentative explanations of any deviations from what is expected. Finally, the conclusion will be associated with the finale of the research project, where the evaluation results have been deemed to be satisficing (“good enough”). Results must be written up, knowledge gained needs to be communicated, and some loose ends (e.g., results that can not properly be explained) may be left for future work (Vaishnavi & Kuechler, n.d.).

As illustrated in Figure 6, knowledge may be created through the distinct phases of development, evaluation, and conclusion, i.e., both through artifact development and through the systematic examination of its use, where the new knowledge should contribute to a refined understanding or contribution to the original problem (related to the environment), and also contribute as new additions to the knowledge base.

DSR proceeds in a cyclical nature (Vaishnavi & Kuechler, n.d.; Hevner, 2007). As such, the DSR process allows for multiple research cycles to be illustrated according to the DSR process model, where one cycle will inform the next. This is in accordance with an adaptation of the Vaishnavi and Kuechler (n.d.) process model that allows for a main cycle of the steps going from problem awareness to conclusion, where the development step may encompass several DSR process subcycles that each contribute new knowledge (van der Merwe et al., 2017).

Complementing the Vaishnavi and Kuechler (n.d.) process model, seven guidelines for effective DSR have been formulated in the influential paper by Hevner et al. (2004). The authors suggest that each guideline should be applied for the research within the DSR framework to be complete. The seven guidelines are: 1) Design as an Artifact; 2) Problem Relevance; 3) Design Evaluation; 4) Research Contributions; 5) Research Rigor; 6) Design as a Search Process; and, 7) Communication of Research. These guidelines have all been followed for the design science research that is part of the research presented in this dissertation.

3.2 Research activities

During this research, two separate methods were applied within the overarching framework of DSR, namely systematic literature review and case study. Here, a systematic review is “a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest" (Kitchenham & Charters, 2007, p. vi). It aims to evaluate research topics in a well-defined, rigorous, and convincing manner (Kitchenham & Charters, 2007). The case study method, in contrast, focuses on one instance (case) related to the subject that is to be investigated, and it takes place in a real-world context (Oates, 2006). A number of data generation methods can be used to gain insights and knowledge about the researched subject, whether qualitative or quantitative, through study of the case (Oates, 2006).

In the following, information will be given regarding the activities conducted as part of this research. The activities are first described more broadly as they pertain to the DSR framework, and then information is provided on the systematic literature review and the case study.

3.2.1 DSR process in the PhD research

This section describes how the activities within the research relate to the design science process model by Vaishnavi and Kuechler (n.d.). Figure 7 shows the DSR process followed and identifies the research activities that more broadly are associated with each step. For the main cycle (Figure 7, column 2), the awareness and suggestion steps were executed in the research that was part of the systematic literature review (Paper 1), while the development step that included two DSR subcycles and the evaluation and conclusion steps were conducted as part of the exploratory case study (documented in Paper 2 and Paper 3).

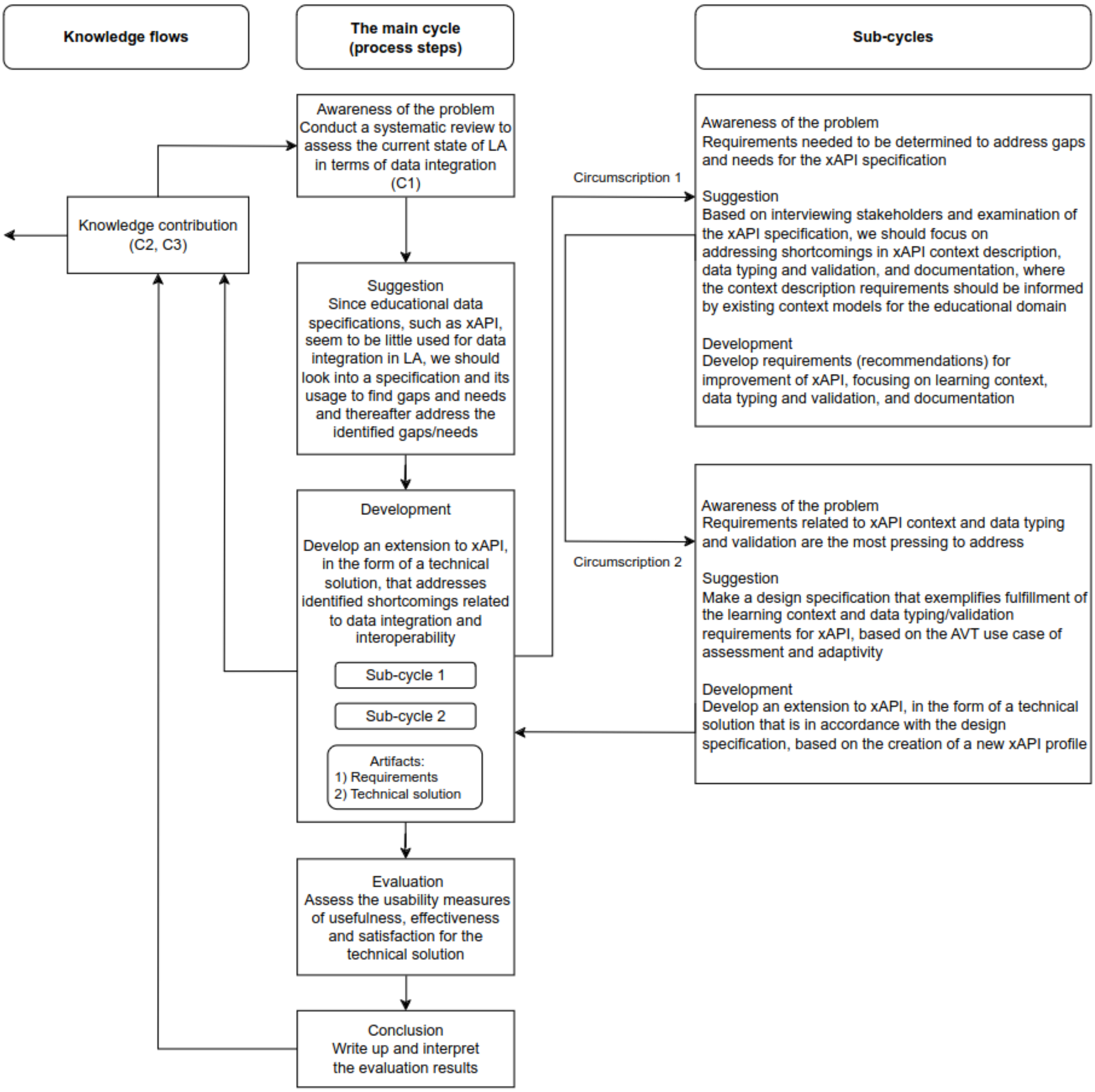

Figure 7 - Design Science Process for PhD Research

As seen in Figure 7, a first awareness of the problem (Figure 7 column 2) arose from conducting a systematic review related to the current state of LA in terms of data integration in the context of higher education (Paper 1). A first understanding of the problem domain was gained through the systematic review, which would enable the identification of gaps and needs in the research literature and provide relevant background knowledge. The suggestion followed from the results and conclusion of the systematic review, which among other findings indicated a lack of use of educational data standards (such as xAPI), thus motivating further examination into the educational data specifications and their usage in order to understand the challenges in adoption and the opportunities afforded from a LA stakeholder perspective.

The development phase of the DSR process was conducted according to two subcycles. The steps of the first subcycle, going from awareness to development, are detailed in the first empirical study (Paper 2). Here, the problem awareness was narrowed to a need to determine requirements to address gaps and needs for the xAPI specification. Based on interviewing LA stakeholders having used xAPI in a real-world setting (i.e., the AVT project), and inspection of the xAPI and xAPI profile specifications, a suggestion was made that the focus should be primarily on addressing gaps and needs for xAPI context, data typing and validation, and documentation. Furthermore, the requirements regarding context should be informed by previous research on context modeling related to the educational domain. The development phase in this subcycle involved the specification of requirements for addressing the gaps and needs in the form of recommendations, informed by the study of the xAPI and xAPI profile specifications and context modeling. The requirements developed in subcycle 1 constituted the first artifact.

The steps of the second development subcycle are documented in the second empirical study (Paper 3). Problem awareness for the second subcycle included a further circumscription, based on the implementable requirements from subcycle 1, to focus more narrowly on xAPI context, data typing and validation. A suggestion was made in the form of a design specification for the recommended changes (requirements), following the affordances and constraints of the xAPI and the xAPI profile specifications, and also taking into account needs derived from the AVT project, i.e., needs to describe concepts related to adaptivity[16]. The development phase in subcycle 2 entailed implementing the design specification in the form of a technical solution based on the creation of a new xAPI profile. Furthermore, an xAPI statement was developed that demonstrated how xAPI statements could adhere to the profile. The

[16] In Paper 3, the design specification is illustrated through Table 1 (column 3), Table 2 and Table 3.

technical solution that implemented the recommendations related to xAPI context, data typing and validation constituted the second DSR artifact, i.e., an instantiation. Finally, the last steps of the DSR process, which were both part of the main DSR cycle, were also documented in the second empirical study (Paper 3). The evaluation step entailed the evaluation of the technical solution according to several usability measures, through user testing with stakeholders (technical experts) from the AVT project who used the technical solution for data description tasks relating to their own or AVT project data. This served to assess if the DSR artifacts actually provided an improvement to the environment. The conclusion step involved writing up and interpreting the evaluation results.

As seen in Figure 7, there is a progression where a problem is first identified, subsequently leading to specification of requirements, and design, before the solution is finally implemented. In this view, the evaluation conducted throughout the research is ex post, i.e., evaluation after actual construction/implementation. It can be argued that this ex post evaluation is naturalistic, since it employs representative users from the AVT project, working with real systems (or standards; i.e., xAPI and the developed technical solution that is implemented based on an xAPI profile), to conduct realistic tasks (assigned by the researchers; stakeholders, however, described their own or AVT project data). While not included in Figure 7, there was also an element of ex ante evaluation throughout the research. For instance, informed arguments were provided by the researchers for the requirements and design relating to the technical solution (some of these were not included in Paper 2/Paper 3 due to journal length restrictions), and additional feedback on requirements for the technical solution was gained from the stakeholders that had real-world experience with xAPI (see Paper 3).

3.2.2 Systematic literature review

A common reason for carrying out a systematic review is to identify gaps in existing research that may inform the direction of future studies. As such, systematic reviews may provide knowledge from which new research activities can be positioned (Kitchenham & Charters, 2007). Thus, a systematic literature review was carried out to assess the current state of data integration in LA focusing on the context of higher education.

The systematic literature review, detailed in Paper 1, was conducted in accordance with the guidelines by Kitchenham and Charters (2007). The systematic review process went through three separate phases: 1) planning the review; 2) conducting the review; and, 3) reporting the review. Review research questions were defined relating to data integration in LA and HE, regarding aspects such as the types of data sources that are used, and how and to what degree different data are being used and combined. A predefined search strategy was determined that aimed to detect the research relevant for the review. Here, a search string was constructed based on the review research questions. Furthermore, academic databases and a number of common venues for publishing learning analytics and educational data mining research were searched. Inclusion and exclusion criteria were used to establish which of the publications returned by the search should be included in the review. The review process was documented using a PRISMA flow diagram (Moher et al., 2009). The data from the included publications were synthesized based on recording of a number of criteria in data extraction forms. Finally, the review was reported (see Paper 1).

3.2.3 Case study

The research conducted and presented in this dissertation has extensively employed the case study method.

The case study research, using a real-world case (i.e., the AVT project) as its subject, investigated challenges and limitations of using a current learning activity data standard (i.e., xAPI) for describing learning context with regard to interoperability and data integration. Additionally, the research investigated how the identified challenges and limitations could be addressed. The case study was exploratory (Oates, 2006, p. 143), meaning the initial activities helped to clarify and understand the topic researched before prompting follow-up activities within the study. Figure 8 gives an overview of the case study research, where the steps above the dotted line are addressed in the research reported in Paper 2, while the steps below the dotted line are addressed in the research reported in Paper 3.

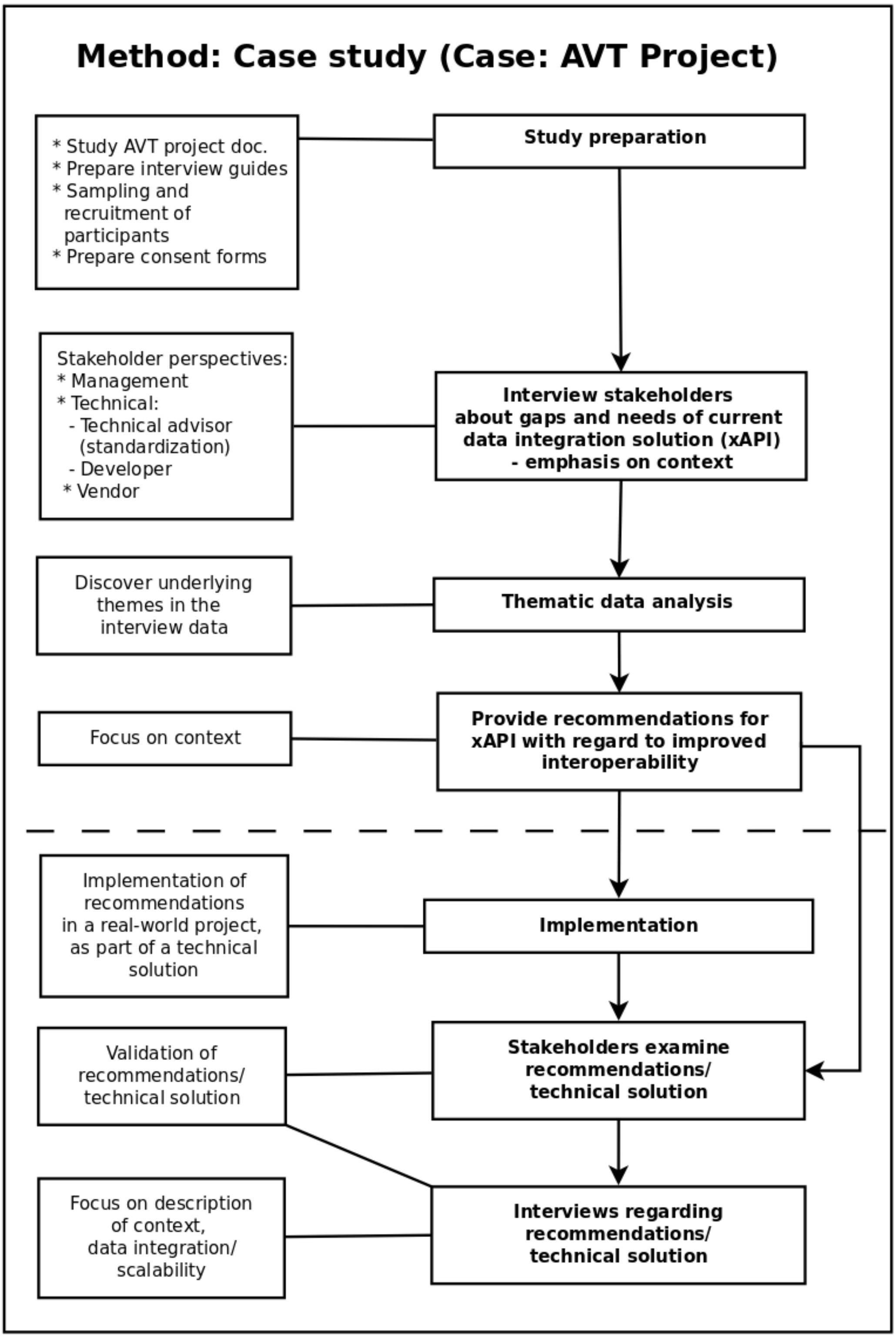

Figure 8 - Case study research overview

Below, a brief explanation related to each of the case study steps (Figure 8) is given. The subsections of this section will subsequently elaborate on different aspects of the case study.

As seen in Figure 8, upon initial study preparation stakeholders from the AVT project were interviewed about the gaps and needs of xAPI, emphasizing context descriptions. After the first set of interviews, and thematic analysis, a set of recommendations to address the identified challenges were established. Next, a technical solution that implemented a number of the recommendations related to xAPI context, data typing and validation, and that was based on the use case of adaptivity, was developed. Finally, the technical solution was validated through user testing, with users performing data description tasks that enabled them to examine the technical solution both at a more general level, related to the recommendations, and at the implemented level. Furthermore, the user testing included interviews to inquire about different aspects of the technical solution.

The case study used a human centered approach (International Organization for Standardization, 2019), which is an approach where real users and their goals influence the development of problem-solving solutions throughout the different stages of a project. Thus, the skill and judgment of real-world users who have domain knowledge through working with xAPI in the AVT project had an influence. Two principles related to the human centered approach which were of special importance for this research are “early focus on users and tasks” and “empirical measurement” (Rogers et al., 2011).

3.2.3.1 The case

The case study investigated the usage of xAPI within the AVT project (Morlandstø et al., 2019; Wasson et al., 2019). The AVT project, a Norwegian research and development project, first ran from June 2017 to May 2019, and then continued under the name AVT2 from October 2019[17]. The project uses a learning activity data standard

[17] In this dissertation, the name “AVT project” is used to describe both the AVT and AVT2 phases of the project.

(i.e., xAPI) to describe activity data regarding Norwegian K-12 students that originate from the EdTech tools/systems of multiple vendors, with the goal to integrate the data across vendors. Due to the project's experience with xAPI data description and data integration it was considered that the project had the potential to reveal challenges and limitations regarding data description and data integration that appear in practice.

The three main partners of the AVT project are the Norwegian Association of Local and Regional Authorities (KS) who own and have funded the project, the Centre for the Science of Learning & Technology (SLATE), University of Bergen who are responsible for research and have led the project, and the Educational Authority in the Municipality of Oslo (Utdanningsetaten). Organizations consulted include the Norwegian Directorate for Education and Training (Utdanningsdirektoratet), the Norwegian Data Protection Authority (Datatilsynet), the Norwegian Competition Authority (Konkurransetilsynet), the Parent organization for schools (Foreldreutvalget for grunnskolen), the Student organization (Elevorganisasjonen), and Standards Norway. Standards Norway, responsible for development and management of Norwegian standards and the Norwegian member organization of the International Organization for Standardization (ISO) and the European Committee for Standardization (CEN), was consulted through the committee SN/K 186[18]. SN/K 186 works more broadly with learning technology and e-learning, and is focused on areas such as learning analytics and metadata (Standards Norway, 2023).

The AVT project explores areas such as learner adaptivity and assessment, where learner adaptivity refers to providing resources (e.g., learning tasks/items) that are better adapted to the learner's needs, and where both formative and summative assessment are of interest to the project (Wasson et al., 2019). As such, reporting and visualizing the activities and progress of learners toward different sets of criteria (e.g., competence objectives) is of importance, in addition to providing learning resource recommendations. The requirements for the system that would contain such functionality were informed by interviews, for example with teachers and vendors

[18] In Norway, the committee is also referred to by the name Læringskomiteen.

(Morlandstø et al., 2019). As mentioned earlier in this section, AVT is concerned with the integration of learning activity data, which provide an improved basis for recommendations and reporting/visualizations. For this purpose, the AVT project examined secure data exchange among vendors, and developed a LA framework that supports the standardization of data originating from different vendor tools and systems. The framework uses xAPI as the standard for data description, integration, and exchange. Aiming for a more consistent use of xAPI, the project has published a vocabulary in an xAPI profile that contains concepts needed for the project, mainly related to the adaptivity use case[19]. In addition, several example xAPI statements for the AVT project, functioning as templates to enable more consistent data descriptions, have also been made publicly available[20]. Several of the vocabulary concepts are based on a vocabulary that had previously been adapted and translated to Norwegian by SN/K 186, and representatives from SN/K 186 also provided input on the xAPI examples.

3.2.3.2 Participants

Participants were involved at several stages of the case study, with each stage having a different number of participants. Participant selection was done using purposive sampling (Bryman, 2012, p. 418), a method that prescribes a strategic way of participant selection, where those selected are considered appropriate based on the study research questions.

In the first instance of user involvement, where participants were interviewed regarding gaps and needs of xAPI, eight participants were recruited. The participants represented different roles related to the AVT project (technical advisor, developer, leader, vendor), and had worked on different tasks. Thus, they offered differing perspectives when interviewed about the gaps and needs of xAPI.

In the second instance of user involvement, user testing that included pre-test and post-test interviews was conducted with three participants to validate the technical solution

[19] AVT focuses on both adaptivity and assessment. A number of existing syntactic structures and vocabulary concepts in xAPI relate to assessment, for instance the xAPI result structure and vocabulary concepts in the SCORM profile. Regarding adaptivity, however, there are fewer existing vocabulary concepts in xAPI. As such, the AVT profile mainly relates to the adaptivity use case.

[20] The profile and examples are available on GitHub: https://github.com/KS-AVT/avt.

that implemented the recommendations related to xAPI data integration and interoperability. The participants were all technical experts that had been affiliated with the AVT project, having experience with xAPI for learning activity data description. Two of the participants had functioned as technical advisors within the AVT project, for instance participating in the development of example xAPI statements and the AVT xAPI profile, and examining how context may be represented using xAPI. They had previously been interviewed to identify gaps and needs of xAPI in the first part of the user involvement. The third participant was a senior developer working for one of the vendors that delivered data to the project.

3.2.3.3 Data collection

Data collection for the case study was conducted first through semi-structured interviews (related to the first instance of user involvement), and subsequently through user testing that included semi-structured interviews (related to the second instance of user involvement).

Semi-structured interview

Interviews are a much used technique for gathering data in case studies (Bryman, 2012, p. 68). The research to inquire about the gaps and needs of xAPI (detailed in Paper 2) used semi-structured interviews for data collection. Relating to their qualitative nature, the semi-structured interviews are influenced by interpretivism in the researcher's understanding of the data (Bryman, 2012).

Eight interviews were conducted with eight participants. The data collection took place primarily from October 22 to November 01, 2019, with a final interview being conducted on February 07, 2020. The first seven of the interviews were conducted face-to-face, while the eighth and final interview was conducted through a digital telecommunications tool. The majority of interviews finished within 60 minutes. All interviews were held in Norwegian.

Three overarching questions were formulated to guide the data collection: 1) *What was the rationale for choosing xAPI in the AVT project?*; 2) *What was the process of using xAPI for data integration and data sharing [in the AVT project]*?; and, 3) *What* *challenges and limitations of xAPI were identified when describing context for data integration?*. Interview guides were developed, supported by a study of AVT project documentation, including the final report for the AVT project (Morlandstø et al., 2019), and inspection of the xAPI and xAPI profile specification. The guide included a list of the main topics that could be covered for the individual interviews, where the topics included: Leadership (e.g., reasons for choosing xAPI for the AVT project); technical development (e.g. experiences of using xAPI for describing data); and, context (information on how the learning context was represented with xAPI and why).

At the beginning of the interviews, the participants were asked to give information on their background, for instance related to the AVT project and experience with xAPI. Since the participants had worked on different roles and tasks, and had different sets of competences, the topics they were asked about from the interview guide would vary. Each topic in the interview guides was, however, covered by at least two of the participants.

Audio was recorded for each of the interviews and the recordings were transcribed. The participants were presented with a consent form before interviews started, all giving consent that data related to their interview may be used for the research presented in Paper 2. In addition, since the audio recordings could in theory be used to identify the participants based on their voice, the study reported in Paper 2 was also reported to Norwegian Centre for Research Data (2020)[21], which approved the study based on the measures taken for privacy and research ethics (see Appendix A). Participant information and audio recordings were stored securely, and no personal identifiable information was registered in the transcripts.

User testing with complementary interviews

User testing involves letting people that are a part of a product's target audience review or try out that product, to evaluate the degree to which the product meets a set of usability measures (Rubin & Chisnell, 2008, p. 21). User testing was conducted to

[21] Starting from 2022, the Norwegian Centre for Research Data has been merged with a number of other organizations into the new organization “Sikt – Norwegian Agency for Shared Services in Education and Research”.

evaluate the usability of the technical solution that implemented the recommendations related to xAPI context, data integration and interoperability, developed as part of this case study (detailed in Paper 3). The overall goal of the user testing was to validate the technical solution in terms of its usefulness, effectiveness and satisfaction, which are related to users' achieving goals/willingness to use a product, alignment with user expectations/simplicity of use, and participant considerations of the product, respectively (Rubin & Chisnell, 2008, p. 4-5). The user test conducted was an assessment test, allowing for evaluation of the solution's lower level aspects and its different aspects, through users performing tasks that mimic those met in the real-world (Rubin & Chisnell, 2008, p. 34-35). Furthermore, semi-structured interviews were conducted to collect data that could complement the data gained from the user test tasks.

Three user tests were conducted with three participants. The data collection took place from November 23, 2022 to December 16, 2022. Data were collected based on pre-test interviews, user test sessions, and post-test interviews. The data from the pre-test interviews were qualitative, regarding participant background. Quantitative performance data, reflecting the participants' behavior (Rubin & Chisnell, 2008, p. 165), were derived based on the user test tasks carried out by the participants; as such, the errors in completion of the user test tasks could be assessed. Finally, the data from the post-test interview were qualitative preference data, reflecting the feelings and opinions of the participants (Rubin & Chisnell, 2008, p. 166) related to the technical solution and its foundations. The pre-test and post-test interviews were conducted through a digital telecommunications tool. All interviews were held in Norwegian.

A number of materials were developed for the user test, including a data registration form, a pre-test interview guide, and a post-test interview guide. The data registration form detailed the user test and described the seven user test tasks that were a part of the user test, where the first six tasks were mandatory and the last task was optional. The tasks, which were related to data descriptions in xAPI, involved filling in different parts of an xAPI statement. All tasks except the first, which tested for general understanding of how to express the xAPI format, focused on filling in information related to xAPI context dimensions and properties, conforming to the xAPI profile developed as part of the technical solution (i.e., the school adaptivity profile). Furthermore, the data registration form included sections to store the xAPI statement that would result from conducting the user test tasks, and to take notes (e.g., regarding thoughts/opinions related to the technical solution). Finally, a number of materials that supported the execution of the user test tasks (e.g., documentation regarding the technical solution) were made available through the data registration form. The pre-test interview guide contained questions about the background of the participants. The post-test interview guide included topics related to the participants' user test experience, the technical solution at a more general level (regarding the recommendations that were implemented), and the implemented solution. The topics covered for example the structuring of data, the context dimensions and properties, and the requirements for context dimensions and properties. The data registration form and interview guides were pilot tested, to allow for identification and correction of any problems with these materials.

A user test began with the pre-test interview, inquiring about the background of the participants. Towards the end of the pre-test interview the participant received information regarding the technical solution that would be tested. Shortly after the pre-test interview ended, participants were emailed the data registration form and an xAPI statement from the AVT project that could be used as a basis for the xAPI statement that would result from the data description tasks (the emailed xAPI statement would need to be modified regarding structure and data types, and furthermore additional data would need to be added; each participant received a separate xAPI statement). Subsequently, the participants could start the user test. Due to the technical expertise of the participants they were asked to go through the user test themselves, thus allowing them to choose when to do the test, and allowing for the participants to conduct the test in their own environment and with their own equipment (e.g., job computer). The participants were encouraged to return the xAPI statement resulting from the user test within a week, by email. Once the participants emailed the xAPI statement resulting from the user test, the post-test interview would be arranged. This interview would inquire about the topics in the post-test interview guide. For this interview the participants could refer to and share the notes that they took during the user test session.

Audio was recorded for each of the pre-test and post-test interviews, with subsequent transcription of the recordings. The participants were presented with a consent form in conjunction with the pre-test interview, and all gave consent that data related to their user test, including pre-test and post-test interview, may be used for the research presented in Paper 3. In addition, since the audio recordings could potentially identify the participants based on their voice, the study reported in Paper 3 was also reported to Norwegian Agency for Shared Services in Education and Research (Sikt) - Data Protection Services[22], which evaluated the processing of personal data to be legal (see Appendix B). Participant information and audio recordings were securely stored, and no personal identifiable information was added to the transcripts.

3.2.3.4 Data analysis

Related to the semi-structured interviews that were detailed in Paper 2, and that were part of the first part of the case study, a thematic analysis approach (Bryman, 2012) was employed to analyze the interview data. The transcripts were collated and read through several times to increase content familiarity. Two levels of coding were applied to the interview data, using the tool NVivo (2022). The data was first coded according to the three overarching questions that would guide the data collection, then subsequently the data coded for each overarching question was coded at a more fine-grained level. Finally, the themes that subsumed a number of codes were identified using an inductive process. The findings for the individual themes were summarized and presented in Paper 2.

Related to the user tests that were detailed in Paper 3, and that were part of the second part of the case study, both data from the user test session and the interview transcripts were analyzed. The post-test interview data analysis was further supported by referring to the participants notes. Regarding the data collected from the user test session, i.e., the xAPI statements resulting from the user test tasks, a quantitative analysis was

[22] https://www.nsd.no/en/data-protection-services

carried out on the completion of the different user test tasks, which each consisted of a number of subtasks. Successful completion of a task required the participant to successfully complete all of its subtasks. Each of the mandatory tasks could at most yield 1 point, provided all of their subtasks were correctly completed, where the subtasks were weighted equally. Consequently, the highest score that could be obtained for each participant was 6 points (the optional task did not count toward the score). The transcribed interview data were collated and read several times, and thereafter analyzed qualitatively. The pre-test interview data were analyzed to obtain participant background information. The post-test interview data were first coded according to questions in the post-test interview guide, which were then mapped to corresponding themes (topics) in the interview guide. Furthermore, data were coded according to themes that had emerged during the interview but were not questions/topics in the interview guide, whereafter these identified themes were aggregated based on an inductive process. The results/findings for the user test, related to user test task completion and the final themes that resulted from the post-test interview analysis, were summarized and presented in Paper 3.

4. Overview of the publications

This chapter presents an overview of the publications that are part of this dissertation. For each paper, the title, author names, and publication venue are included. Additionally, information regarding the following criteria is provided for each paper: research question that is investigated, type of study, methodology, and research outcomes.

4.1 Paper 1

Research paper: Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0105-4

This first paper investigates RQ1: *What are the challenges of current LA approaches and practices for data integration, interoperability, and consequently scalability* to identify challenges relating to the state of the art regarding LA data integration, interoperability, and consequently scalability.

A systematic review enables extensive and rigorous evaluation and interpretation of prior studies that are relevant for a research topic, using a well-defined approach (Kitchenham & Charters, 2007). The goal of the systematic review study detailed in this paper was to assess the current state of data integration in LA, focusing on the higher education context. The need for a systematic review related to learning analytics and data integration was identified first upon doing a more general literature review regarding this research topic, and then further through reading existing reviews and state-of-the-field reports that were relevant to this topic. The focus on higher education was due to the majority of research in learning analytics taking place in higher education (Misiejuk & Wasson, 2017).

The systematic review was conducted according to the guidelines by Kitchenham and Charters (2007). Six academic databases and a number of common venues (journals and conference proceedings) for publishing learning analytics and educational data mining research were searched. The search returned 115 publications, where 64 articles were excluded during an initial screening, and 31 were excluded when assessing the remaining articles for eligibility, thus 20 articles were included in the final analysis. The final articles were published between 2014 and 2019. The final 20 publications were read and reviewed, focusing specifically on a number of criteria (such as "Types of data sources", "Data sources integrated?" and "Data integration approach") that were carefully selected with the aim to understand the LA research with regard to using/combining multiple data sources. Information from each publication was extracted into data extraction forms according to the criteria, and the extracted data were synthesized.

The review reported on the current state with regard to LA and using/combining multiple data sources within higher education, and also enabled the identification of gaps and limitations in existing research on this topic. Additionally, some recommendations were presented to address the identified limitations and gaps.

The literature review revealed that three types of data sources (i.e., LMS, questionnaire, and student information system) appear repeatedly in the research studies, that the number of data sources used/combined in the LA research studies related to the higher education domain tend to be limited (the majority of the studies detailed in the reviewed publications use/combine only 2 data sources), and that when data are integrated they are often in similar formats (which is less of a technical challenge than integrating data from different formats). The review revealed, however, that there seems to have been a shift starting from 2017 to an increasing use of multimodal data and the use/integration of a larger variety of data sources for LA. Furthermore, the findings indicated that publications that report on LA research tend to lack details about integration of data in the implemented systems, and that there is a lack of stakeholder involvement (e.g., teacher, EdTech vendor) in the LA research studies. A gap identified is that while educational data specifications such as xAPI and Caliper Analytics may be a good starting point for data integration, especially as they relate to the technical level of interoperability and have capabilities for the semantic level, they do not seem to be widely used. This prompted a recommendation for further investigation into such standards, including opportunities and challenges for their adoption and use, focusing on the LA stakeholder point of view.

The identification of challenges relating to the state of the art regarding data integration, interoperability, and consequently scalability for LA, specifically the indicated lack of use of educational data specifications despite such specifications being a promising approach for data integration, informed the direction of the further research (Paper 2 and Paper 3).

4.2 Paper 2

Research paper: Samuelsen, J., Chen, W., & Wasson, B. (2021). Enriching context descriptions for enhanced LA scalability: a case study. *Research and Practice in Technology Enhanced Learning, 16*(1). https://doi.org/10.1186/s41039-021-00150-2

This second paper investigates RQ1: *What are the challenges of current LA approaches and practices for data integration, interoperability, and consequently scalability?* to identify gaps and needs regarding data integration and interoperability, emphasizing learning context descriptions, for a learning activity data standard (i.e., xAPI). Furthermore, the paper examines RQ2: *How can descriptions of context in a current learning activity data standard, i.e., xAPI, be enhanced in order to provide improved data integration, interoperability, and consequently scalability?* to identify how context descriptions in a learning activity data standard (i.e., xAPI) can be enriched for improved data integration and interoperability.

This paper details the first empirical study. In this study, based on the previous finding regarding the indicated lack of use of educational data specifications, the aim was to understand and address challenges of using learning activity data standards for data integration and interoperability. The learning activity data standard examined was xAPI, and the focus was to a large degree on descriptions of the learning context. This study constituted the first part of an exploratory case study, using the AVT project as the case. While the AVT project involves schools (K-12) and not higher education, it had used xAPI for describing data originating from different data sources, and could therefore be expected to uncover challenges of data description and data integration as they unfold in a real-world setting. These challenges will be relevant regardless of the educational sector in which xAPI is applied.

To elicit gaps and needs for xAPI with a focus on learning context, semi-structured interviews were conducted with 8 stakeholders from the AVT project. The stakeholders represented different perspectives (e.g., leader, developer, technical advisor) and had worked on different tasks within AVT. Three overarching questions were formulated to guide the data collection: 1*) What was the rationale for choosing xAPI in the AVT project?*; 2) *What was the process of using xAPI for data integration and data sharing [in the AVT project]*?; and, 3*) What challenges and limitations of xAPI were identified when describing context for data integration?*. The interview data were analyzed using a thematic analysis approach (Bryman, 2012).

The thematic analysis enabled the identification of underlying themes in the data. Following an inductive process, seven themes emerged, where each theme pertained to one of the overarching questions that guided the data collection. Related to the question regarding the rationale for choosing xAPI, two themes were identified: "Open, flexible, and mature specification" and "familiarity". Related to the question regarding the process of using xAPI for data integration and data sharing, two themes emerged: "Practical experimentation and learning by trial and error" and "expert-driven technical process". Finally, related to the question regarding the challenges and limitations of xAPI for data integration when describing contextual data, three themes emerged: "Data description constructs", "(semantic) differences between tools generating statements", and "semantic vs. technical interoperability".

From the thematic analysis, it was found that there is a lack of clarity in how to describe contextual data in xAPI, which may create barriers to data integration and interoperability. Based on findings from the analysis, and inspection of the xAPI and xAPI profile specifications, a number of challenges for xAPI and data integration were identified, pertaining primarily to context descriptions, but also to data typing and validation, and documentation. This resulted in the development of a set of recommendations, provided in Paper 2 in summarized form, for how to address the identified challenges. The recommendations were based on a search of the knowledge base for relevant literature on context categorizations in the educational domain, and were informed by affordances and constraints available in the xAPI and xAPI specifications.

One of the main recommendations was to use a hierarchical (2-level) and unified structure for xAPI context, consisting of context dimensions with adhering properties, to promote clarity in how to describe context in xAPI, and thus also expressibility. Further use of data typing and validation was also among the recommendations, for instance to clarify and enforce the assumptions of xAPI, and to make explicit the expected level of data granularity for different tool developers. These recommendations were at the conceptual level, and while a number of them could be implemented through developing and utilizing an xAPI profile, it was acknowledged that some recommendations would require a standard change in order to be implemented.

The recommendations that were part of this paper were expressed at the conceptual level. As such, this paper informed the direction of the third paper, where the recommendations that had been given at the conceptual level were examined as to how they could be implemented as part of a technical solution.

4.3 Paper 3

Research paper: Paper 3: Samuelsen, J., Chen, W., & Wasson, B. (submitted). Implementing enriched context descriptions for efficient scaling of Learning Analytics. *Journal of Learning Analytics*.

The paper examines RQ2: *How can descriptions of context in a current learning activity data standard, i.e., xAPI, be enhanced in order to provide improved data* *integration, interoperability, and consequently scalability?* to identify how context descriptions in a learning activity data standard (i.e., xAPI) can be enriched for improved data integration and interoperability, focusing on implementation as part of a technical solution. Furthermore, the paper examines RQ3: *How can the technical solution for enhanced context descriptions in xAPI be validated?* to investigate how the technical solution can be validated.

This paper details the second empirical study, constituting the second part of the exploratory case study using the AVT project as the case. The aim of this paper was twofold: The first aim was to implement the recommendations regarding xAPI data integration and interoperability, focusing on context, as part of a technical solution. The second aim was to validate the technical solution through user testing with technical experts that have previous experience with data description using xAPI.

First the recommendations were explored as to how they might translate to a more general technical solution, applying more broadly to the xAPI and xAPI profile specifications, and, subsequently, how they could be implemented as part of a technical solution. Here, implementation of the recommendations as part of a technical solution was based on a study of the xAPI and xAPI profile specifications, considering possibilities and constraints. Furthermore, the technical solution was adapted for adaptivity, a main focus area for the AVT project.

The implemented technical solution was enabled through creation of an xAPI profile, i.e., the school adaptivity profile. With the profile, recommendations could be expressed, such as the unified and hierarchical expression of context through context dimensions with adhering properties. First, a number of properties previously identified as important by teachers (Morlandstø et al., 2019; Hansen et al., 2020) were classified according to context dimensions. Second, some further context dimensions were defined, with adhering properties, acknowledging that learning may happen outside the physical boundary of the school and where the context dimensions had been indicated as important for adaptive learning by previous research (Thüs et al, 2014; Lincke, 2020). Finally, a number of properties that were a part of the xAPI context structure, and contextual information that had previously been modeled using contextActivities, were reclassified according to context dimensions. In total, 6 context dimensions with adhering properties were modeled (see Paper 3). Among the xAPI profile constructs utilized for the school adaptivity profile is ContextExtension, where JSON schema definitions enabled the definition of context dimensions/properties and specification of requirements for the properties (e.g., data typing and validation measures).

To validate the implemented recommendations, the technical solution went through user testing with 3 technical experts from the AVT project. The overall objective of user testing was to assess the usability measures usefulness, effectiveness, and satisfaction for the solution. Data were collected from the user testing session where the participants went through a number of data description tasks, and furthermore from pre-test and post-test interviews. The data from the user testing sessions, i.e., the resulting xAPI statements, were assessed for successful completion of the user test tasks according to quantitative measures, allowing the identification of errors in task completion. The pre-test interview data were analyzed to identify participant background information, while the post-test interview data were analyzed qualitatively according to topics.

The results from the user testing sessions showed that the participants had been able to successfully complete the majority of the mandatory six user test tasks and their sub-tasks (the participants achieved scores of 5.83, 5.83, and 5.33 out of 6 points).

The findings from the post-test interviews were categorized according to nine topics identified during the analysis. The topics stemmed from the post-test interview guide and the interview data.

In general, the interview findings indicated that the user test instructions were considered understandable, as confirmed by the test session results. Furthermore, the findings revealed that the participants preferred the unified and hierarchical structuring of contextual data to the current xAPI approach of using a flat structure where data may be spread across different structures within the context. The findings also indicated that the unified and hierarchical structuring is more readable than the current solution, and adds to aspects such as cleaner organization, learnability, and understandability, thus promoting expressibility. While expressibility may promote data integration, pragmatic views were expressed related to data integration, i.e., that post hoc processing and clear guidelines could be other alternatives for achieving integration.

Regarding the context dimensions and properties, four of the six context dimensions were found to be appropriate to include based on the focus on adaptivity in K-12 education, while there was uncertainty regarding inclusion of the final two context dimensions. Regarding the requirements for context dimensions and properties for the technical solution, it was indicated that data typing and validation measures can promote data integration, while it was simultaneously cautioned that such measures need to fit reality. Some conditions were identified where there may be a discrepancy between the (technical solution's) model constraints and real-world constraints. Regarding the implementation of the technical solution that was implemented through the creation of a new xAPI profile, comments were given on a number of functionality such as statement templates, the ContextExtension construct and the JSON schema specification that was utilized by the construct to represent context dimensions and constraints for the context dimensions and properties. The participants suggested that use of the ContextExtension construct with the JSON schema specification provided an interesting way to extend the basic capabilities of xAPI, but that other schema languages than JSON schema (e.g., XML Schema Definition [XSD]) may be more expressive. Considering inclusion of the JSON schemas in the profile (inline) or through definitions hosted using IRIs, there was a preference for utilizing external schemas, although it was suggested that validation tools may have better support for inline schemas. Furthermore, the findings indicated that the xAPI specification, and thus the technical solution, could benefit from introducing linked data technologies, for example by expressing xAPI statements that may adhere to a profile using JSON-LD rather than JSON.

Some potential improvements were suggested for the technical solution, for example regarding context dimension and property naming, properties that could be added/removed, and context dimensions and property data typing and validation measures.

Related to the overall user test goal of assessing the usability measures usefulness, effectiveness, and satisfaction for the technical solution, the user test indicated that the usefulness and effectiveness criteria were met. Furthermore, satisfaction was expressed regarding aspects of the general technical solution, while some room for improvement was expressed for the implemented solution.

5. Discussion

In this chapter, the research questions are first evaluated, related to how they have been addressed through the individual contributions. Subsequently, related research is examined. Next, methodological considerations are discussed, related to design science research framework and the research approach. Finally, limitations of the research are presented.

5.1 Evaluation of research questions

To fulfill the objectives of this research, three research questions were formulated. Within the broader framework of design science research, this research has utilized systematic literature review (related to RQ1) and exploratory case study methodologies (related to RQ1, RQ2 and RQ3) as approaches to answer the research questions.

In the following, the research questions are evaluated regarding how they have been addressed through the contributions.

RQ1: What are the challenges of current LA approaches and practices for data integration, interoperability, and consequently scalability?

This research question is answered through the contributions C1 and C2. Through systematically reviewing the literature, a number of challenges were identified related to LA core issues of scalability (C1), including that there is little use of educational data specifications even though they are a promising approach for data integration. To further examine educational data standards and their use, and for understanding challenges of educational data standards adoption for LA, gaps were identified related to xAPI expressibility (C2). In addition to context descriptions, the identified gaps were related to other aspects such as data typing and validation, which all negatively impact xAPI expressibility.

RQ2: How can descriptions of context in a current learning activity data standard, i.e., xAPI, be enhanced in order to provide improved data integration, interoperability, and consequently scalability?

This research question is answered through the contributions C2 and C3. A conceptual solution was developed in this research that consists of recommendations to enhance the expressibility of xAPI, focusing on xAPI context, data typing and validation (C2). The search for the recommendations was based on both the study of xAPI and previous research on context modeling for the educational domain. Among the recommendations are: 1) a unified and hierarchical (2-level) representation of context, using context dimensions with adhering properties; and, 2) further use of data typing and validation, for example to restrict the properties and values for context dimensions, and to support correct use of xAPI based on its assumptions.

Furthermore, a technical solution was developed in this research that implements the recommendations, based on the K-12 adaptivity case that is of importance for the AVT project (C3). The technical solution was implemented through the creation of an xAPI profile for the K-12 adaptivity domain, i.e., the school adaptivity profile, taking into account the affordances and constraints of the xAPI and xAPI profile specifications.

RQ3: How can the technical solution for enhanced context descriptions in xAPI be validated?

This research question is answered through contribution C3. The technical solution that implements the recommendations for enhanced expressibility of xAPI, focusing on context, was subject to evaluation. As such, both the technical solution and the recommendations that functioned as a basis for the technical solution were assessed. The approach chosen for evaluation was to employ user testing, where assessment test design enabled usability evaluation regarding the solution's low-level aspects through users performing authentic tasks. Technical experts affiliated with the AVT project were given a number of user test tasks related to data description in accordance with the technical solution that implements the recommendations. Furthermore, a post-test interview inquired about different aspects of the technical solution and, at a more conceptual level, the recommendations that were implemented through the technical solution. The technical solution was assessed according to the usability criteria of usefulness, effectiveness, and satisfaction.

5.2 Related work

In this section, related research concerning approaches for context modeling in the educational domain, the xAPI data model and representation of xAPI context, and the LA architecture approach to data integration will be discussed.

5.2.1 Approaches for context modeling

Verbert et al. (2012) surveyed different context elements, for example context dimensions and properties, that are used as part of recommender systems in the educational domain. Furthermore, a number of studies look at the expression of context related to learning (e.g., Jovanović et al., 2007; Schmitz et al., 2011; Thüs et al., 2012; Lincke, 2020). The developed context models have varied in aspects such as their unit of focus (e.g., learner versus learning object), flexibility regarding data registration, and categorization of context (e.g., flat structure versus hierarchical structure).

The research on expression of context influenced our recommendation for modeling context in xAPI, where context is modeled according to a hierarchical and unified structure, and additional measures for data typing and validation are recommended to enhance xAPI expressibility. However, the context models reviewed have not previously been applied to xAPI, and they are not part of de facto learning activity data standards that are used by industry, as such they do not constitute viable alternatives to xAPI.

5.2.2 Extensions of the xAPI data model and xAPI context

Some research has been conducted related to representing and extending the xAPI data model. Research has looked at representing the general xAPI data model using ontologies (e.g., Vidal et al., 2015; De Nies et al., 2015; Vidal et al., 2018). Furthermore, different research has focused on extending xAPI for use with particular use cases or communities of practice (e.g., Anseeuw et al., 2016; Scheffel et al., 2016; Bakharia et al., 2016; Serrano-Laguna et al., 2017; Nouira et al., 2019; Keehn & Claggett, 2019). Some of the related works have acknowledged the need for collecting contextual data, and as such extended xAPI with contextual concepts using xAPI profiles or recipes[23], for example for self-regulated learning, systemic analytics and serious games (Scheffel et al., 2016; Bakharia et al., 2016; Di Mitri et al., 2017; Keehn & Claggett, 2019). In addition, other related works have introduced new contextual properties to the xAPI data model using ontologies (Anseeuw et al., 2016; Nouira et al., 2019). However, none of these works have focused on representing context in a unified and hierarchical way. Moreover, these other works are related to other use cases and/or communities of practice (such as self-regulated learning, serious games, and assessment analytics) than this research that targets the K-12 school adaptivity domain. This research defines a number of context dimension properties that have been indicated as useful through interviews with teachers (Hansen et al., 2020). These properties have been categorized according to a solution for hierarchical classification based on context dimensions, with some further context dimensions and properties being added based on previous research indicating their importance (Thüs et al, 2014; Lincke, 2020). Here, the context dimensions also encompass properties that have previously been spread across different structures within the xAPI context, to accommodate a unified context representation. Additionally, this research not only focuses on context modeling, but also data typing and validation as a means to improve xAPI expressibility.

5.2.3 LA architectures

LA architectures, such as the Apereo, JISC, and SURFnet architectures, provide approaches for integration of data originating from multiple data sources (Sclater, 2017). The integration of data, which is supported through utilization of for example plugins/connectors and data transformations, may be enabled by representing the collected data according to a common standard, for example xAPI or Caliper. As such, learning activity data can be stored collectively in a data store such as an LRS, according to a common data model and format. Furthermore, learning activity data

[23] The xAPI recipe is an older concept than the xAPI profile concept. It contains textual descriptions on how to represent certain types of learning activities using xAPI (Bakharia et al., 2016; Miller, 2018). While recipes and profiles may be created for the same type of purpose, recipes do not offer the machine-readability afforded by xAPI profiles.

stored in an architecture may be formatted according to profiles or recipes (Berg, Mol et al., 2016).

While LA architectures provide a step toward integration of learning activity data, they do not address the underlying challenges of the learning activity data standards. For xAPI such underlying challenges include that the standard allows for the same type of data to be stored according to different context structures, and that the openness and flexibility of xAPI allow the same types of data to be represented in a myriad of ways. Furthermore, because of the openness and flexibility of xAPI, different architectures that support the xAPI standard can represent the same type of activity data differently. As such, the problem of data integration remains when combining learning activity data that originate from different architectures.

In contrast, this research has looked at the expressibility of data descriptions of xAPI, i.e., the consistent description of data across sources, focusing on context, data typing and validation measures. Such an approach complements the LA architecture approaches, seeking to enable a more consistent approach to xAPI data descriptions, thus also facilitating data integration across LA architectures that utilize the standard.

5.3 Methodological considerations

This section provides a discussion regarding how the design science research framework was applied for this research and an evaluation of the research approach.

5.3.1 Design science research

This research has applied design science research (DSR) as the overarching research framework. As stated in the Methodology chapter, the seven DSR guidelines by Hevner et al. (2004) were followed for the research. How the individual guidelines have been addressed for the research that is presented in this dissertation will be discussed in the following.

Design as an artifact

This guideline states that DSR must result in viable artifacts (Hevner et al., 2004). The artifacts should be purposeful, addressing unsolved real-world problems. The artifact types listed by Hevner et al. (2004) are construct, model, method, or instantiation. More recently, however, Offermann et al. (2010) have made additions to the DSR artifact typology, including requirements and patterns.

This research has resulted in two artifacts. The first artifact is a set of requirements, in the form of recommendations to improve the expressibility of xAPI, focusing on xAPI context. The second artifact is an instantiation where the requirements that constitute the first artifact have been implemented as part of a technical solution.

Problem relevance

This guideline states that DSR aims to build/develop solutions to important organizational problems (Hevner et al., 2004).

The artifacts have been developed as part of an exploratory case study, using the AVT project as the case. As such, they address real and unsolved problems regarding data integration and LA scalability that apply to organizations such as educational institutions and EdTech vendors.

Design evaluation

This guideline states that “the utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods” (Hevner et al., 2004, p. 83).

The artifacts were subject to evaluation as part of the case study. The technical solution that implemented the requirements was validated through user testing with technical experts from the AVT project, which assessed the usability criteria of usefulness, effectiveness, and satisfaction. The user test was an assessment test, where the participants performed tasks that mimic those met in the real-world, and included complementary pre-test and post-test interviews. As such, the participants could explore and reason about both the recommendations that were implemented through the technical solution and its various functionality. Since the evaluation employed real users (technical experts) from the AVT project, working with real systems (or standards; i.e., xAPI and the technical solution that is implemented through the development of the school adaptivity profile), and the tasks performed are realistic in that they mimic those conducted in the real-world, it can be argued that the evaluation is naturalistic.

Research contributions

This guideline states that clear design contributions must be provided as a result of effective DSR. The design of artifacts may lead to three types of contributions: The artifacts that were designed, design foundations, and design methodologies (Hevner et al., 2004).

This design science research has resulted in two artifacts. The first artifact is the requirements (recommendations) to improve xAPI with regard to expressibility, focusing on the learning context. The second artifact is the technical solution, which implements the requirements based on the AVT use case of K-12 adaptivity. Furthermore, the creation of both artifacts was motivated by challenges identified through conducting the systematic review to assess the current state of data integration in LA, related to core issues of LA scalability.

Research rigor

This guideline states that both the building and evaluation of design science research artifacts needs to apply rigorous methods. The researcher needs to carefully select and use appropriate theories and methods from the knowledge base (Hevner et al., 2004).

Research rigor was addressed through the careful application of theory and methods within the exploratory case study. Semi-structured interviews were conducted to identify gaps and needs of xAPI data integration, focusing on expressibility and context, with the interview data being subject to thematic analysis. The gaps and needs identified informed the recommendations developed. Furthermore, the technical solution that implemented the recommendations was rigorously evaluated through user testing, which included complementary interviews. Finally, the artifacts constructed as part of this research were also informed by theory and techniques in the knowledge base, which included literature on xAPI and on context data modeling related to education.

Design as a search process

This guideline states that “the search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment” (Hevner et al., 2004, p. 83). Since it is typically infeasible to find an optimal solution, heuristic search is applied to build satisficing solutions that work well in organizational environments for the problems specified (Hevner et al., 2004).

Heuristic search strategies were applied during the design science research, to aid in identifying possible design alternatives that could be iteratively tested to see if they fit the requirements and constraints for the artifacts that were to be built. Furthermore, the scope of the design science research was iteratively expanded, as progress was made on different parts of the research (as illustrated through the DSR subcycles).

Communication of research

This guideline states that the design science research should be effectively communicated to audiences both within technology and the management traditions (Hevner et al., 2004).

The design science research, including the systematic review that informed the construction of the artifacts, has been communicated in 3 journal papers (where two papers have been published, and the third has been submitted for review). Furthermore, the research is presented in this dissertation. When communicating the research related to construction and evaluation of the artifacts, details have been provided that are mainly relevant for a more technology-oriented audience, such as practitioners and researchers. Furthermore, the knowledge communicated related to the application of the artifacts, for example which problems they address and the usefulness of the solutions, can help managerial audiences to decide if they should expend resources on implementing/adapting the presented solutions to their organizational context.

A number of technical presentations of the work have been given during the research period, including to the working group ISO/IEC JTC 1/SC 36/WG 8 "Learning Analytics Interoperability"[24] and to the Standards Norway SN/K 186 "Learning Technology" committee, which both have audiences from industry (e.g., EdTech vendors) and academia. The research has also been presented to representatives from the GIGA project in Japan, which focuses on among others learning activity data interoperability for all Japanese K-12 schools.

5.3.2 Research approach

Reliability and validity are often seen as important when evaluating social science research (Bryman, 2012). Reliability regards if the results of a study can be repeated, and is concerned with "issues of consistency of measures" (Bryman, 2012, p. 168). Validity typically relates to accuracy of measurements and if the research measures that which it should measure (Golafshani, 2003). Researchers have argued that these criteria are better suited for quantitative research within the positivist research paradigm than qualitative research, for example qualitative interviews, which relates to the interpretivist paradigm (Golafshani, 2003). As interpretivism operates on the assumption that reality is a subjective construct that differs among individuals and groups, it can be argued that there is no single truth within interpretivist research. Furthermore, qualitative researchers do not aim to detach themselves from or manipulate the research settings, but rather stay involved and immersed in the research that unfolds in a real-world setting, for example having interactions with the research participants through interviews, meaning that a researcher can have some effects on both the research participants and the situation being studied, which can again affect the results. Therefore, it is unlikely that one researcher will obtain the exact same results as another researcher. As such, the reliability criteria of repeatability is likely to be unattainable within qualitative research (Golafshani, 2003; Oates, 2006). Lincoln and Guba (1985) suggest dependability, relating to how well the research process and

[24] https://www.iso.org/committee/45392.html

data have been described/documented, as an alternative criteria to reliability for qualitative research.

The term of validity, as it is typically used in the positivist tradition, has been considered as inadequate by researchers in qualitative research, leading qualitative researchers to either redefine the term or search for alternative terms (Golafshani, 2003). Lincoln and Guba (1985) specify trustworthiness as an alternative to validity. Here, trustworthiness concerns the amount of trust that can be granted to specific research (Oates, 2006). One aspect related to trustworthiness is credibility, which concerns if the research was carried out in a means that ensured accurate identification and descriptions of the subject of study, leading to credible findings. Measures to enhance credibility include researcher engagement with the research problem for an extended period of time and triangulation (e.g., utilizing multiple data sources or methods). Another aspect related to trustworthiness is transferability, relating to if findings in one case may be transferable to other cases. Here, it is useful to give detailed descriptions regarding the context of the study (Oates, 2006; Bryman, 2012).

For this research, the dependability of the research was sought through detailed descriptions of the research process, including data collection, as expressed in the dissertation chapter 3 and chapter 4, and the individual papers. The aim has been to be as transparent as possible. Furthermore, this research has aimed for trustworthiness, seeking to address both credibility and transferability criteria. Credibility has been sought through prolonged exposure to the research challenges, and triangulation. Regarding triangulation, the systematic literature review detailed in Paper 1 utilized two researchers to analyze and judge the relevance of the literature identified for the review, which is in accordance with the guidelines by Kitchenham and Charters (2007), constituting a type of investigator triangulation (Oates, 2006). Furthermore, the user test detailed in Paper 3 utilized both assessment testing and qualitative interviews as data generation methods, to assess usability criteria for the technical solution and to explore and reason about the recommendations that were implemented through the technical solution and its various functionality. Finally, regarding the transferability of the research, extended descriptions of the research case have been provided, so that others may consider the transferability of the research, for example its relevance to their own situation.

5.4 Limitations

This research has several limitations that may be considered. First, the systematic literature review detailed in Paper 1 focused on LA data integration in higher education. As such, some potential research that integrates data, taking place in other educational levels, may not have been taken into account when assessing the state of data integration in LA. Second, due to the qualitative nature of the research detailed in Paper 2, it can be argued that the research findings do not generalize. Rather, qualitative research is utilized in an attempt to clarify, understand, and extrapolate to cases that are similar (Golafshani, 2003). Many of the challenges identified, however, have support in the existing research literature, as stated in Paper 2. Third, the interviews and user testing conducted as part of this research had only a limited number of participants, all recruited from the AVT project as part of the case study. Eight AVT stakeholders participated in the interviews, while three stakeholders participated in the user testing. While the interviews recruited stakeholders having different perspectives (e.g., leader, vendor, technical advisor), the usability testing recruited technical experts. Here, two of the three technical experts that user tested the technical solution were technical advisors that had previously been interviewed regarding the gaps and needs for xAP that informed the recommendations that were implemented in the technical solution. Finally, the context dimensions and the classification of properties according to the dimensions as well as the requirements for the context dimension properties (e.g., format and data type) have not yet been validated by domain experts working within K-12 education, such as teachers. Validating these by domain experts will be able to provide further insights into the applicability and usability of the technical solution.

6. Conclusion and future work

Data integration, i.e., the collection and combination of data originating from multiple sources, is closely linked to interoperability and is an important aspect that contributes to Learning Analytics (LA) scalability. This PhD research has sought to investigate challenges of LA data integration, interoperability, and thus scalability, focusing on identifying gaps and needs for a learning activity data standard, i.e., xAPI. Furthermore, solutions have been contributed to the identified gaps and needs.

A systematic review was first conducted to examine the state of LA data integration, allowing for the identification of challenges in combining data sources. Based on the review, it was identified that learning activity data standards do not seem to be widely used for data integration in LA, prompting inquiry into why there seems to be limited standards use. An exploratory case study approach was followed, using the Activity Data for Assessment and Adaptivity (AVT) project as the case. AVT was chosen as the case due to its experience with utilizing xAPI for data descriptions, aiming to integrate K-12 learner data originating from multiple data sources. The case study consisted of two parts. In the first part of the case study, challenges and limitations of xAPI were examined, focusing on context and expressibility. The examination was based on in-depth interviews with stakeholders from the AVT project, and furthermore on the inspection of xAPI. Based on thematic analysis of the interview data, and inspection of xAPI, recommendations were provided, at the conceptual level, for how xAPI may be improved with regard to expressibility, focusing on learning context. Among the recommendations was to utilize a hierarchical and unified structure for context descriptions, consisting of context dimensions with adhering properties. In the second part of the case study, the recommendations were implemented as part of a technical solution, following affordances and constraints of the xAPI and xAPI profile specifications and based on the K-12 school adaptivity use case that is of importance for the AVT project. Furthermore, the technical solution was evaluated by means of user testing, including complementary interviews, with technical experts. This allowed the participants to explore and assess the solution both at a more general level, related to the recommendations, and at the implemented level.

Some future work is presented as follows. First, the validation was conducted with only a small number of technical experts, and while potential improvements were identified during the validation of the technical solution, the technical solution could be validated by a greater number of technical experts to identify further room for improvement. Second, teachers and other domain experts for the K-12 domain could inspect the context dimensions, the classification of properties according to the dimensions, and the property requirements that are all part of the implemented technical solution, to ensure they make sense with regard to K-12 and adaptivity. The technical solution could be refined based on feedback from the experts during the user testing, for instance in terms of modifications of the context dimensions and properties. Future work could also look into cross-standard interoperability, i.e., how to ensure that xAPI may interoperate with other standards. Here, linked data technologies may provide a basis for exploration, as they provide a generic way to model the data using the RDF data model, allow for mapping of terms across standards communities, and may pertain both to data structures and data values. Finally, aspects of the technical solution, including the unified and hierarchical structuring of context descriptions, should be further explored by the xAPI community for inclusion in the standard as a means to promote scalable LA.

References

1EdTech (n.d.-a). Caliper Analytics. Retrieved October 13, 2022, from https://www.imsglobal.org/activity/caliper

1EdTech (n.d.-b). Initial xAPI/Caliper Comparison. Retrieved October 13, 2022, from https://www.imsglobal.org/initial-xapicaliper-comparison

Advanced Distributed Learning. (2017a). xAPI specification - part one: About the experience API. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-About.md#partone

Advanced Distributed Learning. (2017b). xAPI specification - part three: Data processing, validation, and security. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-Communication.md#partthree

Advanced Distributed Learning. (2020). Anticipating the xAPI Version 2.0 Standard. Retrieved from https://adlnet.gov/news/2020/08/06/Anticipating-the-xAPI-Version-2.0-Standard/

Advanced Distributed Learning. (2021a). Toward the Harmonization of Learning Activity Metadata. Retrieved from https://www.adlnet.gov/news/2021/05/28/P2881-and-the-Harmonization-of-Learning-Metadata/

Advanced Distributed Learning. (2021b). xAPI specification. Retrieved from https://github.com/adlnet/xAPI-Spec

Advanced Distributed Learning. (2021c). xAPI Specification - Part Two: Experience API Data. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-Data.md#parttwo

Advanced Distributed Learning. (2022a). xAPI profiles specification. Retrieved from https://github.com/adlnet/xapi-profiles

Advanced Distributed Learning. (2022b). xAPI profile specification - part two: xAPI profiles document structure specification. Retrieved from https://github.com/adlnet/xapi-profiles/blob/master/xapi-profiles-structure.md#part-two

Allemang, D., Hendler, J., & Gandon, F. (2020). *Semantic Web for the Working Ontologist*. Elsevier.

Anseeuw, J., Verstichel, S., Ongenae, F., Lagatie, R., Venant, S., & De Turck, F. (2016). An ontology-enabled context-aware learning record store compatible with the experience api. In *8th International joint conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KEOD)* (pp. 88-95).

Apereo. (2020). Learning Analytics Initiative | Apereo. Retrieved September 1, 2020 from https://www.apereo.org/communities/learning-analytics-initiative

Baca, M. (Ed.). (2016). *Introduction to metadata*. Getty Publications.

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016). Recipe for success: Lessons learnt from using xAPI within the connected learning analytics toolkit. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 378–382).

Bailey, W. (2005) What is adl scorm. Retrieved from http://publications.cetis.org.uk/wp-content/uploads/2011/10/WhatIsScorm2_web.pdf

Benson, T., & Grieve, G. (2016). Why Interoperability Is Hard. In T. Benson & G. Grieve, *Principles of Health Interoperability* (pp. 19–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-30370-3_2

Berg, A. M., Mol, S. T., Kismihók, G., & Sclater, N. (2016). The role of a reference synthetic data generator within the field of learning analytics. *Journal of Learning Analytics*, *3*(1), 107-128.

Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016). Dutch cooking with xAPI recipes: The good, the bad, and the consistent. In *2016 IEEE 16th international conference on advanced learning technologies (ICALT)* (pp. 234-236). IEEE.

Berners-Lee, T., & Fischetti, M. (1999). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. Harper San Francisco.

Betts, B., & Smith, R. (2019). The learning technology manager's guide to xAPI (Version 2.2). Retrieved from https://learningpool.com/guide-to-xapi/

Bizer, C., Heath, T., & Berners-Lee, T. (2011). Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts* (pp. 205-227). IGI global.

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford: Oxford University Press.

Buckingham Shum, S. (2012). Learning analytics. *UNESCO Institute for Information Technologies in Education*. Retrieved from http://iite.unesco.org/files/policy_briefs/pdf/en/learning_analytics.pdf

Buckingham Shum, S., & McKay, T. (2018). Architecting for learning analytics: Innovating for sustainable impact. *EDUCAUSE Review*.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5/6), 318-331. https://doi.org/10.1504/IJTEL.2012.051815

Chatti, M. A., Lukarov, V., Thüs, H., Muslim, A., Yousef, A.M. F., Wahid, U., Greven, C., Chakrabati, A., & Schroeder, U. (2014). Learning Analytics: Challenges and Future Research Directions. *E-Learning and Education, 10*(1).

Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an open learning analytics ecosystem. In B. K. Daniel (Ed.), *Big Data and learning analytics in higher education* (pp. 195–219) https://doi.org/10.1007/978-3-319-06520-5_12.

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134-138).

Cooper, A. (2013). Learning Analytics Interoperability-a survey of current literature and candidate standards.

Cooper, A. (2014a). Learning analytics interoperability-the big picture in brief. *Learning Analytics Community Exchange*.

Cooper, A. (2014b). Specifications and Standards - Quick Reference Guide. *Learning Analytics Community Exchange*.

Cooper, A., & Hoel, T. (2015). Data sharing requirements and roadmap. Public Deliverable D7.2 from the LACE project.

De Meester, B., Lieber, S., Dimou, A., & Verborgh, R. (2018). Interoperable user tracking logs using {linked data} for improved learning analytics. In *Call Your DATA Proceedings of the XIXth international CALL research conference* (pp. 29-31).

Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). *Learning pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data* (pp. 188–197). ACM Press https://doi.org/10.1145/3027385.3027447.

De Nies, T., Salliau, F., Verborgh, R., Mannens, E., & Van de Walle, R. (2015). TinCan2PROV: exposing interoperable provenance of learning processes through experience API logs. In *Proceedings of the 24th international conference on World Wide Web* (pp. 689-694).

Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing*, *5*(1), 4–7.

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). *Rethinking learning analytics adoption through complexity leadership theory* (pp. 236–244). ACM Press https://doi.org/10.1145/3170358.3170375.

Duval, E. (2004). Learning Technology Standardization: Making Sense of it All. *Computer Science and Information Systems*, *1*(1), 33-43.

Duval, E. (2012) Learning Analytics and Educational Data Mining. Retrieved from http://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/data-mining/

European Commission. (2017). New European Interoperability Framework. Retrieved from https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf

European Union. (2016). Regulations. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

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

Flanagan, B., & Ogata, H. (2017). Integration of learning analytics research and production systems while protecting privacy. In *The 25th International Conference on Computers in Education, Christchurch, New Zealand* (pp. 333–338).

Flanagan, B., & Ogata, H. (2018). Learning analytics platform in higher education in Japan. *Knowledge Management and E-Learning*, *10*(4), 469–484.

Friesen, N. (2005). Interoperability and learning objects: An overview of e-learning standardization. *Interdisciplinary Journal of E-Learning and Learning Objects*, *1*(1), 23-31.

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, *48*, 108-119.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report*, *8*(4), 597-607.

Gordon, J., Hayden, T., Johnson, A., & Smith, B. (2020). *Total learning architecture 2019 report*. Advanced Distributed Learning Initiative. https://adlnet.gov/publications/2020/04/2019-Total-Learning-Architecture-Report/

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, *15*(3), 42-57.

Griffiths, D., & Hoel, T. (2016). *Comparing xAPI and Caliper* (Learning Analytics Review, No. 7). Bolton: LACE.

Griffiths, D., Hoel, T., & Cooper, A. (2016). Learning analytics interoperability: Requirements, specifications and adoption. Public Deliverable D7.4 from the LACE project.

Hakimi, L., Eynon, R., & Murphy, V. A. (2021). The ethics of using digital trace data in education: A thematic review of the research landscape. *Review of educational research*, *91*(5), 671-717.

Hansen, C.J.S., Wasson, B., & Belokrys, G. (2020). *Teacher's need for information.* Internal report: unpublished.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian journal of information systems*, *19*(2), 87-92.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75-105.

IEEE. (2022) P 92741.1 xAPI Base Standard. Retrieved from https://sagroups.ieee.org/9274-1-1/

International Organization for Standardization. (2019). Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (ISO Standard No. 9241-210:2019).

International Organization for Standardization. (2020). Information technology for learning, education and training — Learning analytics interoperability — Part 3: Guidelines for data interoperability (ISO Standard No. 20748-3:2020).

Jain, P., Gyanchandani, M., & Khare, N. (2016). Big data privacy: a technological perspective and review. *Journal of Big Data*, *3*(1), 1-25.

JISC. (2020). Learning records warehouse: technical overview: Integration overview. Retrieved September 01, 2020 from https://docs.analytics.alpha.jisc.ac.uk/docs/learning-records-warehouse/Technical-Overview:%2D%2DIntegration-Overview

Jovanović, J., Gašević, D., Knight, C., & Richards, G. (2007). Ontologies for effective use of context in e-learning settings. *Journal of Educational Technology & Society*, *10*(3), 47–59.

Khalil, M., & Ebner, M. (2015). Learning analytics: Principles and constraints. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Canada (pp. 1326–1336).

Keehn, S., & Claggett, S. (2019). Collecting standardized assessment data in games. *Journal of Applied Testing Technology*, *20*(1), 43–51.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, Keele University (UK).

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Kitchin, R., & Lauriault, T. P. (2015). Small data in the era of big data. *GeoJournal*, *80*(4), 463-475.

Kitto, K., Whitmer, J., Silvers, A. E., & Webb, M. (2020). Creating Data for Learning Analytics Ecosystems. *Solar Position Paper*.

Kivunja, C., & Kuyini, A. B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education*, *6*(5), 26-41. https://doi.org/10.5430/ijhe.v6n5p26

Lincke, A. (2020). A Computational Approach for Modelling Context across Different Application Domains (PhD dissertation, Linnaeus University Press). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-93251

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Sage.

Lukarov, V., Chatti, M. A., Thüs, H., Kia, F. S., Muslim, A., Greven, C., & Schroeder, U. (2014). Data Models in Learning Analytics. In *DeLFI Workshops* (Vol. 1014, pp. 88-95).

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision support systems*, *15*(4), 251-266.

Miller, B. (2018). Profile Recipes vs. xAPI Profiles [Blog post]. Retrieved from https://xapi.com/blog/profile-recipes-vs-xapi-profiles/

Miller, P., & Duan, X. (2018). NGDLE Learning Analytics: Gaining a 360-Degree View of Learning. Retrieved from https://er.educause.edu/blogs/2018/1/ngdle-learning-analytics-gaining-a-360-degree-view-of-learning

Misiejuk, K., & Wasson, B. (2017). *State of the Field Report on Learning Analytics. SLATE Report 2017-2*. Centre for the Science of Learning & Technology (SLATE), University of Bergen.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA group. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*.

Morlandstø, N. I., Hansen, C. J. S., Wasson, B., & Bull, S. (2019). *Aktivitetsdata for vurdering og tilpasning: Sluttrapport. SLATE Research Report 2019-1*. Bergen: Centre for the Science of Learning & Technology (SLATE). ISBN: 978-82-994238-7-8.

Muslim, A. (2018). *OpenLAP: a user-centered open learning analytics platform* (Doctoral dissertation, Dissertation, RWTH Aachen University, 2018).

Norwegian Centre for Research Data (2020). NSD - Norwegian Centre for Research Data. Retrieved August 20, 2020, from https://nsd.no/nsd/english/index.html

Nouira, A., Cheniti-Belcadhi, L., & Braham, R. (2019). An ontology-based framework of assessment analytics for massive learning. *Computer Applications in Engineering Education*, *27*(6), 1343-1360.

NVivo. (2022) Qualitative Data Analysis Software | NVivo. Retrieved from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Oates, B. J. (2006). *Researching information systems and computing*. SAGE Publications.

Ochoa, X., & Worsley, M. (2016). Augmenting learning analytics with multimodal sensory data. *Journal of Learning Analytics*, *3*(2), 213-219.

Offermann, P., Blom, S., Schönherr, M., & Bub, U. (2010). Artifact types in information systems design science: A literature review. In R. Winter, J. L.Zhao, & S. Aier (Eds.), *Global Perspectives on Design Science Research: DESRIST 2010* (pp. 77-92). Springer.

Pomerantz, J. (2015). *Metadata*. MIT Press.

Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). Strategies for design science research evaluation.

Rabelo, T., Lama, M., Vidal, J. C., & Amorim, R. (2017). Comparative study of xAPI validation tools. In *2017 IEEE Frontiers in Education Conference (FIE)* (pp. 1-5). IEEE.

Rogers, Y., Sharp, H. & Preece, J. (2011). Interaction Design: beyond human-computer interaction (3rd ed.). Wiley.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons.

Scheffel, M., Ternier, S., & Drachsler, H. (2016). *The Dutch xAPI Specification for Learning Activities (DSLA) – Registry*. Retrieved from http://bit.ly/DutchXAPIreg

Schmitz, H. C., Wolpers, M., Kirschenmann, U., & Niemann, K. (2011). Contextualized attention metadata. In *Human attention in digital environments* (pp. 186–209).

Sclater, N. (2017). *Learning analytics explained*. Routledge.

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education: A review of UK and international practice. Retrieved from https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v3.pdf

Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, *50*, 116-123.

Shankar, S. K., Prieto, L. P., Rodríguez-Triana, M. J., & Ruiz-Calleja, A. (2018). A review of multimodal learning analytics architectures. In *2018 IEEE 18th international conference on advanced learning technologies (ICALT)* (pp. 212-214). IEEE.

Shankar, S. K., Rodríguez-Triana, M. J., Prieto, L. P., Calleja, A. R., & Chejara, P. (2022). CDM4MMLA: Contextualized data model for multimodal learning analytics. *In the Multimodal Learning Analytics Handbook*. Springer.

Siemens, G. (2011). 1st international conference on learning analytics and knowledge. Technology Enhanced Knowledge Research Institute (TEKRI). Retrieved from https://tekri.athabascau.ca/analytics/

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*, 30–32. https://doi.org/10.1145/2330601.2330605

Siemens, G., & Baker, R. S. D. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254).

Smith, B., & Milham, L. (2021). Total Learning Architecture (TLA) Data Pillars and Their Applicability to Adaptive Instructional Systems. In *International Conference on Human-Computer Interaction* (pp. 90-106). Springer, Cham.

Sonnenberg, C., & vom Brocke, J. V. (2011). Evaluation patterns for design science research artefacts. In *European design science symposium* (pp. 71-83). Springer, Berlin, Heidelberg.

Sonnenberg, C., & vom Brocke, J. V. (2012). Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. In *International Conference on Design Science Research in Information Systems* (pp. 381-397). Springer, Berlin, Heidelberg.

Standards Norway. (2023). SN/K 186. Retrieved from https://www.standard.no/standardisering/komiteer/sn/snk-186/

Thüs, H., Chatti, M. A., Brandt, R., & Schroeder, U. (2015). Evolution of interests in the learning context data model. In *Design for Teaching and Learning in a Networked World*, (pp. 479–484). Cham: Springer.

Thüs, H., Chatti, M. A., Greven, C., & Schroeder, U. (2014). Kontexterfassung,-modellierung und-auswertung in Lernumgebungen. *DeLFI 2014-Die 12. e-Learning Fachtagung Informatik.*

Thüs, H., Chatti, M. A., Yalcin, E., Pallasch, C., Kyryliuk, B., Mageramov, T., & Schroeder, U. (2012). Mobile learning in context. *International Journal of Technology Enhanced Learning*, *4*(5-6), 332–344.

Vaishnavi, V., Kuechler, B. (n.d.). Design science research in information systems. Retrieved from: http://www.desrist.org/design-research-in-information-systems/

van der Merwe, A., Gerber, A., & Smuts, H. (2017). Mapping a design science research cycle to the postgraduate research report. In *Annual Conference of the Southern African Computer Lecturers' Association* (pp. 293-308). Springer, Cham.

van der Merwe, A., Gerber, A., & Smuts, H. (2020). Guidelines for Conducting Design Science Research in Information Systems. In B. Tait, J. Kroeze, & S. Gruner (Eds.), *ICT Education* (Vol. 1136, pp. 163–178). Springer International Publishing. https://doi.org/10.1007/978-3-030-35629-3_11

Venable, J. (2006). A framework for design science research activities. In *Emerging Trends and Challenges in Information Technology Management: Proceedings of the 2006 Information Resource Management Association Conference* (pp. 184-187). Idea Group Publishing.

Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., & Duval, E. (2012). Context-aware recommender systems for learning: a survey and future challenges. *IEEE transactions on learning technologies*, *5*(4), 318-335.

Vidal, J. C., Rabelo, T., & Lama, M. (2015). Semantic description of the Experience API specification. In *2015 IEEE 15th International Conference on Advanced Learning Technologies*, (pp. 268–269).

Vidal, J. C., Rabelo, T., Lama, M., & Amorim, R. (2018). Ontology-based approach for the validation and conformance testing of xAPI events. *Knowledge-Based Systems*, *155*, 22-34.

Wasson, B., Morlandstø, N. I., & Hansen, C. J. S. (2019). *Summary of SLATE Research Report 2019-1: Activity data for assessment and activity (AVT)*. Bergen: Centre for the Science of Learning & Technology (SLATE). Retrieved from https://bora.uib.no/handle/1956/20187

Winne, P. H. (2017). Learning analytics for self-regulated learning. In L. Charles, S. George, W. Alyssa, & G. Dragan (Eds.), *Handbook of learning analytics* (pp. 241–249). Beaumont: Society for Learning Analytics Research.

Yin, R. K. (2014). *Case study research: design and methods* (5th ed.). SAGE Publications.

Zouaq, A., Jovanovic, J., Joksimovic, S., & Gasevic, D. (2017). Linked data for learning analytics: Potentials and challenges. In L. Charles, S. George, W. Alyssa, & G. Dragan (Eds.). *Handbook of Learning Analytics* (pp. 347-355). Beaumont: Society for Learning Analytics Research.

Appendices

A. Approval by the Norwegian Centre for Research Data (NSD) for study reported in P2

Sikt

Meldeskjema / Enriching LA context descriptions for enhanced scalability: A case st… / Vurdering

Vurdering av behandling av personopplysninger

Referansenummer	Vurderingstype	Dato
561403	Standard	07.09.2020

Prosjekttittel
Enriching LA context descriptions for enhanced scalability: A case study

Behandlingsansvarlig institusjon
Universitetet i Bergen / Det psykologiske fakultet / SLATE-Centre for the Science of Learning & Technology

Prosjektansvarlig
Jeanette Samuelsen

Prosjektperiode
01.04.2019 - 14.09.2020

Kategorier personopplysninger
Alminnelige

Lovlig grunnlag
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene er lovlig så fremt den gjennomføres som oppgitt i meldeskjemaet. Det lovlige grunnlaget gjelder til 14.09.2020.

Meldeskjema

Kommentar
Bekreftelse på status

NSD har vurdert endringen registrert 28.08.2020.

Vi har nå registrert 14.09.2020 som ny sluttdato for forskningsperioden.

NSD vil følge opp ved ny planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.
Lykke til videre med prosjektet!

Kontaktperson hos NSD: Elizabeth Blomstervik
Tlf. personverntjenester: 55 58 21 17 (tast 1)

B. Approval by "Sikt - Data Protection Services" for study reported in P3

Sikt

Meldeskjema / Implementing enriched contextual descriptions for efficient scaling of L… / Vurdering

Vurdering av behandling av personopplysninger

Referansenummer	Vurderingstype	Dato
541529	Standard	20.12.2022

Prosjekttittel
Implementing enriched contextual descriptions for efficient scaling of Learning Analytics(arbeidstittel)

Behandlingsansvarlig institusjon
Universitetet i Bergen / Det psykologiske fakultet / SLATE-Centre for the Science of Learning & Technology

Prosjektansvarlig
Jeanette Samuelsen

Prosjektperiode
03.01.2022 - 31.01.2023

Kategorier personopplysninger
Alminnelige

Lovlig grunnlag
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene er lovlig så fremt den gjennomføres som oppgitt i meldeskjemaet. Det lovlige grunnlaget gjelder til 31.01.2023.

Meldeskjema

Kommentar
OM VURDERINGEN
Personverntjenester har en avtale med institusjonen du forsker eller studerer ved. Denne avtalen innebærer at vi skal gi deg råd slik at behandlingen av personopplysninger i prosjektet ditt er lovlig etter personvernregelverket.

Personverntjenester har nå vurdert den planlagte behandlingen av personopplysninger. Vår vurdering er at behandlingen er lovlig, hvis den gjennomføres slik den er beskrevet i meldeskjemaet med dialog og vedlegg.

VIKTIG INFORMASJON TIL DEG
Du må lagre, sende og sikre dataene i tråd med retningslinjene til din institusjon. Dette betyr at du må bruke leverandører for spørreskjema, skylagring, videosamtale o.l. som institusjonen din har avtale med. Vi gir generelle råd rundt dette, men det er institusjonens egne retningslinjer for informasjonssikkerhet som gjelder.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til den datoen som er oppgitt i meldeskjemaet.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
Personverntjenester vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen

formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet

lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), og dataportabilitet (art. 20).

Personverntjenester vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
Personverntjenester legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Ved bruk av databehandler (spørreskjemaleverandør, skylagring eller videosamtale) må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29. Bruk leverandører som din institusjon har avtale med.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til oss ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde: https://www.nsd.no/personverntjenester/fylle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema

Du må vente på svar fra oss før endringen gjennomføres.

OPPFØLGING AV PROSJEKTET
Personverntjenester vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Part 2
The papers

Paper 1

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0105-4

Samuelsen et al. Research and Practice in Technology Enhanced Learning (2019) 14:11
https://doi.org/10.1186/s41039-019-0105-4

Research and Practice in Technology Enhanced Learning

RESEARCH

Open Access

Integrating multiple data sources for learning analytics—review of literature

Jeanette Samuelsen[1,2*], Weiqin Chen[1,3] and Barbara Wasson[1,2]

* Correspondence: Jeanette.Samuelsen@uib.no
[1]Centre for the Science of Learning & Technology, University of Bergen, Bergen, Norway
[2]Department of Information Science & Media Studies, University of Bergen, Bergen, Norway
Full list of author information is available at the end of the article

Abstract

Learning analytics (LA) promises understanding and optimization of learning and learning environments. To enable richer insights regarding questions related to learning and education, LA solutions should be able to integrate data coming from many different data sources, which may be stored in different formats and have varying levels of structure. Data integration also plays a role for the scalability of LA, an important challenge in itself. The objective of this review is to assess the current state of LA in terms of data integration in the context of higher education. The initial search of six academic databases and common venues for publishing LA research resulted in 115 publications, out of which 20 were included in the final analysis. The results show that a few data sources (e.g., LMS) appear repeatedly in the research studies; the number of data sources used in LA studies in higher education tends to be limited; when data are integrated, similar data formats are often combined (a low-hanging fruit in terms of technical challenges); the research literature tends to lack details about data integration in the implemented systems; and, despite being a good starting point for data integration, educational data specifications (e.g., xAPI) seem to seldom be used. In addition, the results indicate a lack of stakeholder (e.g., teachers/instructors, technology vendors) involvement in the research studies. The review concludes by offering recommendations to address limitations and gaps in the research reported in the literature.

Keywords: Learning analytics, Higher education, Data integration, Multiple data sources, Interoperability, Scalability

Introduction

Learning analytics (LA) includes collecting, computationally analyzing, and reporting data to stakeholders, to gain insights and enable decision-making and interventions related to questions about learning and learning environments (Siemens, 2011). The data can be stored in different formats, have varying levels of structure, and originate from different sources. When data from multiple sources are collected and merged, i.e., the data are integrated, they may reflect the dispersed activities of learners in a more precise way than what is possible for each individual data source (Chatti, Muslim, & Schroeder, 2017). In addition, data integration can also lead to more useful analysis, since many LA techniques require large-scale and possibly diverse data (Cooper & Hoel, 2015). Data integration is one contributing factor to the scalability of LA. Factors important for scaling up LA include the technical solution but also factors such as organizational hierarchy

(Buckingham Shum & McKay, 2018), management structures (Dawson et al., 2018), policy and regulations (European Union, 2016).

Data integration is closely related to interoperability, which involves technical, semantic, legal, and organizational levels. The *semantic level* is about ensuring that data format and meaning is preserved and understood. The *technical* level includes services for data exchange and data integration. The *legal level* addresses aspects such as enabling collaboration despite different legal frameworks and organizational policies. With regard to this interoperability level, an important factor is to protect the privacy of users. Finally, *organizational* interoperability includes aligning processes for common organizational goals and addressing user expectations and requirements (European Commission, 2017).

Data integration and interoperability are general challenges when working with Big Data and are not specific to any one domain (Kadadi, Agrawal, Nyamful, & Atiq, 2014). In an educational setting, collecting and combining data from multiple sources can help provide insights that have implications for areas such as learning, instruction, retention, and curriculum design. The data collected are often activity data (i.e., data from learners' activity in digital learning environments). As stated by Chatti et al. (2017), adopting widely used specifications is important for interoperability. Within education and LA, two well-known educational data specifications are xAPI (2019) and IMS Caliper Analytics (2019). These two specifications enable the exchange of data among different applications, and integration of data from multiple sources in a central data store. In addition, the specifications provide a specific format and syntax to describe learning events that have occurred in learning environments; and, they enable (but do not enforce) the use of controlled vocabularies, which can be used to define concepts and relationships for describing and representing a domain of interest (Allemang & Hendler, 2011). Thus, the two specifications target the technical and, at least to some degree, semantic interoperability levels.

In most cases, only a limited number of data sources are used in combination for data analysis in the context of LA in higher education. Existing LA projects addressing data integration needs tend to place emphasis on the technical level of interoperability (Apereo, 2018; JISC, 2019; OnTask, 2019). The new EU general data protection regulation (European Union, 2016) addresses, to a large degree, legal and organizational concerns. Semantic interoperability, enabling shared data meaning, is typically less emphasized, even though it can enable more effective merging of data through reuse of common data specifications. While all the interoperability levels are important for data integration, the focus in this review is on the semantic and technical levels.

In this article, we report on a systematic literature review that examines publications on LA research studies in higher education that use and/or combine data from multiple sources. The reason for focusing on higher education as the domain of study is twofold: (1) higher education is the working context of the authors and the datasets we are using in our studies come from a higher education setting and (2) previous research has found that the majority of LA research is conducted in higher education (Misiejuk & Wasson, 2017); thus, this restriction of scope would not be expected to significantly affect the number of publications included in the review.

The aim of our research is to assess the current state of LA in terms of data integration in the context of higher education. As a conclusion, we identify shortcomings in existing research and provide recommendations to address such limitations and gaps.

Review method and results

This systematic review follows the guidelines by Kitchenham and Charters (2007). The following review stages will be detailed: planning the review, conducting the review, and data synthesis and reporting the results.

Planning the review

In the planning phase, we clarified the needs for a review, by going through relevant existing reviews and state-of-the-field reports (e.g., Misiejuk & Wasson, 2017; Sclater, Peasgood, & Mullan, 2016; Shahiri, Husain, & Rashid, 2015) and identified gaps in the knowledge base. This process resulted in the following questions that this review seeks to answer:

- *RQ1*. What types of data and data sources are being used for LA in higher education?
- *RQ2*. How and to what extent are different data being used/combined for LA research in higher education?
- *RQ3*. What methods are used, and what issues are being addressed through using/combining multiple data sources?

Conducting the review

Six academic databases were initially selected for search: ACM Digital Library, IEEE Xplore, SpringerLink, Science Direct, Wiley, and AISEL. Later, we added common venues for publishing LA and the related field of educational data mining (EDM) research as additional sources: Journal of Learning Analytics (jLA), Journal of Educational Data Mining, International Conference on Learning Analytics & Knowledge, International Conference on Educational Data Mining, and Learning at Scale (L@S) conference. For journals, issues from 2014 (when jLA was first published) and later were included. For conferences, we included proceedings from 2017, 2018 and 2019 (the most recent conferences, as research in earlier conferences would most likely have been published in journals by this time).

The search string was constructed based on the research questions, as follows:

("multiple data sources" OR multimodal OR "multi-modal" OR "multiple data sets" OR "multiple datasets") AND ("learning analytics" OR "educational data mining") AND "higher education"

As can be seen, the search string encompasses relevant fields (LA and EDM), context (higher education), and concepts (multiple data sources).

Before beginning the systematic literature review, we defined inclusion and exclusion criteria. These criteria were based on the research questions and helped with the assessment of the relevance of each publication for the review. The inclusion and exclusion criteria are listed in Tables 1 and 2 below.

Conducting the review included the following steps: identification, screening, eligibility, and inclusion. Figure 1 shows the results of each step using a PRISMA flow diagram (Moher, Liberati, Tetzlaff, Altman, & PRISMA group, 2009).

The initial search of academic databases was conducted on September 5, 2017. Forty-nine results (journal articles) were returned from the academic databases, no duplicates were among the results. Two researchers judged the relevance of the publications for the

Table 1 Inclusion criteria

Inclusion criteria	Explanation
Publication domain	The domain of the study must be higher education or massive open online courses (MOOCs).
Publication type	Studies must be peer-reviewed and published in a journal/conference proceedings.
Data source	The research must make use of multiple data sources.
Implementation	Publications must include frameworks for combining data sources or details about implemented systems using/combining multiple data sources (not just proposed design and/or architecture).
Language	Publications must be written in English.

systematic review in two phases. First, the title, abstract, and keywords were read for each publication, to filter out results that were not relevant for the review based on the inclusion and exclusion criteria. This initial screening resulted in 26 publications being excluded. Second, the remaining 23 results were assessed for eligibility, based on reading the full-text of the publications and checking this against the inclusion and exclusion criteria. This phase removed 17 publications and resulted in six for inclusion in the review. On April 18, 2019, to include publications that had been published since the initial search, an identical search string was applied to search the same academic databases. This search yielded 32 new results (journal articles). During the initial screening, 19 publications were excluded. The remaining 13 publications were assessed for eligibility, leading to the removal of eight publications. Thus, five new publications were included in the review following this newer search.

Additional literature was downloaded from online repositories for common venues for publishing LA and EDM research, between July 24 and July 26, 2018. The downloaded literature was indexed based on full-text PDF files, thereafter the review search string was applied on the indexed literature. Indexing and search were done using Zotero (reference management software). Initially, the search string resulted in 22 publications (journal articles and conference papers). Assessing the relevance of the publications for the systematic review, 11 publications were removed during the initial screening; then, three publications were removed during full-text assessment for eligibility. Thus, eight additional publications were found eligible for inclusion. On April 18, 2019, we added all new research from the LA and EDM venues to Zotero (i.e., new conference proceedings and journal issues), before re-applying the full-text search. This search gave 12 new results. Out of the resulting publications, eight were removed during the initial screening, and three were removed when assessing the publications for eligibility. Thus, this newer search of LA and EDM venues resulted in one new publication being included in the systematic review.

In total, there are 20 publications included in the review. They are listed in Table 3.

Table 2 Exclusion criteria

Exclusion criteria	Explanation
Data analysis	Publications that primarily focus on data analysis methods and algorithms, but have little information regarding data, are excluded.
Publication type	Literature reviews are excluded.

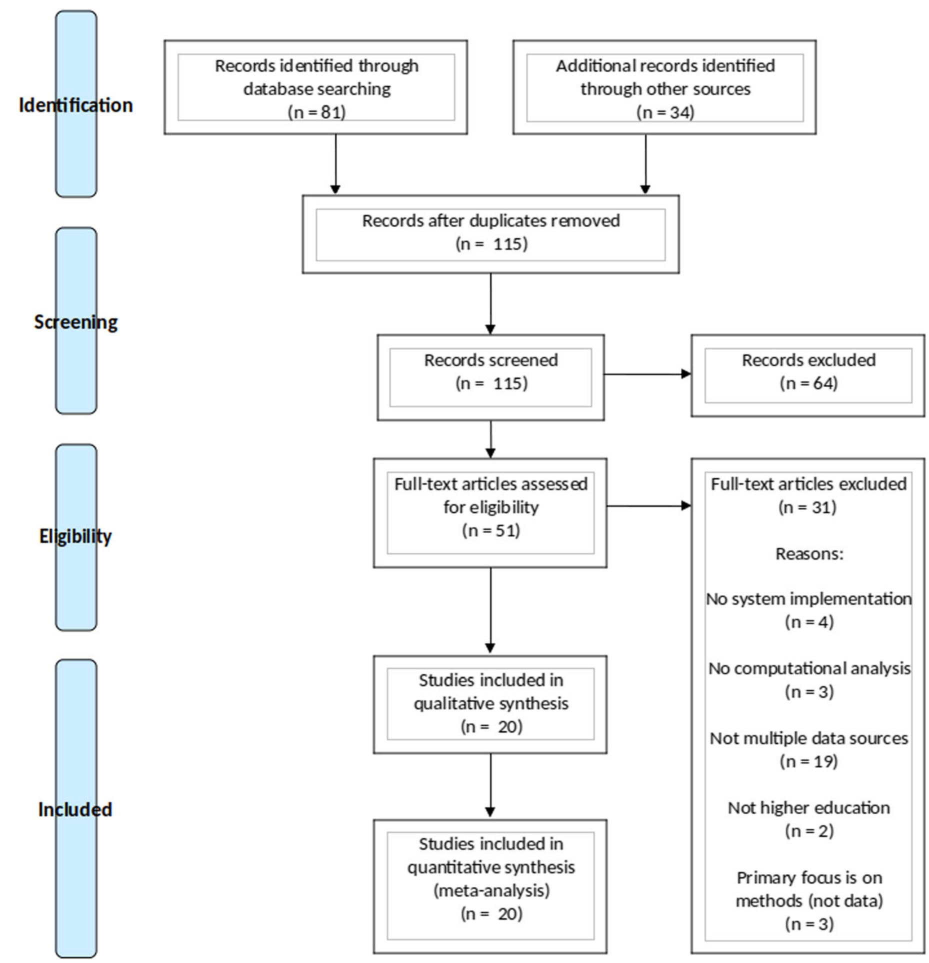

Fig. 1 Review process diagram

Data synthesis and reporting the results

The publications included in the review were read, reviewed (using data extraction forms to record details for every relevant criterion) and synthesized. The following criteria were assessed for each publication (see analysis of the articles in Table 3):

- Types of data sources
- Types of data
- Data sources integrated?
- Data integration approach (manual or automatic)
- System supports educational data specifications?
- Issue addressed in publication
- Methods used
- Records for how many participants analyzed
- Data storage technology (see Table 7)

Integration of data sources was judged based on the information in the publications' text about integration, or (if the former was not available) which data were analyzed in combination to address a specific issue. Similarly, we looked for clues

Table 3 Overview of the reviewed publications

Publication	Issue addressed in publication	Types of data sources	Types of data	Data sources integrated?	Data integration approach (manual or automatic)	Methods used	Records for how many participants analyzed
Lopez Guarin, Guzman, and Gonzalez (2015)	Predict the loss of academic status at a certain time	Student information system (× 2)	Student background information (× 2), performance test data, final grades	Yes	Automatic—first join admissions data sets into one table. Then join with academic information	Decision trees, naive Bayes	1532 students
Park, Yu, and Jo (2016)	Classify blended learning courses in a Korean higher education institution	LMS (× 2)	Activity log, course data	Yes	Automatic (most likely—not explicitly stated). Combine course data and log data on course ID (anonymized)	Latent class analysis	N/A (Records regarding 612 courses which were found suitable for analysis)
Thompson, Kennedy-Clark, Wheeler, and Kelly (2014)	Automatic tagging of text part of speech; for the identification of types of micro-events that learners enact, and the determination of whether learners complete functions that are crucial for task success	Corpora (× 2) (both are mini-corpora of collaborative problem-based learning activities)	Text (× 2)	No (data are analyzed separately)	Data are not integrated	Part of speech tagger—trained on Penn Tree Bank corpora. Visualization of timing and speaker for each utterance in one mini corpora.	Corpora 1: total 6 dyads (12 students + teacher) Corpora 2: four postgraduate students
Zheng, Bender, and Nadershahi (2017)	Data were extracted from tools to provide data on faculty's application of digital tools and to assess the impact of the lecture annotation tool on students' learning behavior	LMS, lecture annotation tool	Activity log data (× 2)	No (data are analyzed separately)	Data are not integrated	N/A	N/A

Table 3 Overview of the reviewed publications *(Continued)*

Publication	Issue addressed in publication	Types of data sources	Types of data	Data sources integrated?	Data integration approach (manual or automatic)	Methods used	Records for how many participants analyzed
Pardos and Kao (2015)	Bayesian network analysis to assess student current and prior knowledge for problems in a MOOC (with visualization); and visualization of course structure (not based on preceding analysis)	MOOC (× 2)	Activity log, student background information (possibly more)	No, platform can currently only integrate EdX MOOC data with other EdX MOOC data. Platform also supports Coursera	Automatic for integrating EdX MOOC data with other EdX data (for Coursera MOOC data this is not addressed). Approach: use HarvardX tool to integrate different types of EdX files into one csv file (loosely based on xAPI). For visualizations: read csv file(s) into memory	Bayesian network analysis, visualization	N/A
Liu et al. (2017)	Examine use of an adaptive system through analysis of usage patterns	Student information system, adaptive platform, LMS, performance test	Student background information, activity log, performance test data (× 3)	Yes	N/A (publication explicitly mentions combination of data, yet does not specify how)	Spearman correlation, visualizations, regression analyses	128 first-year students entered into pharmacy program
Raca, Tormey, and Dillenbourg (2016)	Compare student behaviors (levels of movement) and connect with attention (self-reported)	Video, questionnaire	Video-derived data, questionnaire data	Yes	N/A	Descriptive statistics (e.g., mean, percentage), correlations	56 bachelor level students
Di Mitri et al. (2017)	Predict learners performance during self-regulated learning	Physiological signals wristband, software tracking tool, questionnaire, weather information	Physiological arousal data, software category, questionnaire data, location data, weather data	Yes	Automatic. A tool (Learning Pulse Server) imports data from different APIs and stores events in a Learning Record Store (xAPI format)	Linear mixed effects models	9 PhD students (the multimodal data set originally contained approximately 10,000 records)
Ochoa et al. (2018)	Provide automatic feedback on oral presentation skills	Video, audio, presentation slide	Video derived data, audio derived data, presentation slide derived data	No (data are analyzed separately)	Data/data sources are not integrated	Various classification algorithms (e.g., random forest)	83 engineering students

Table 3 Overview of the reviewed publications *(Continued)*

Publication	Issue addressed in publication	Types of data sources	Types of data	Data sources integrated?	Data integration approach (manual or automatic)	Methods used	Records for how many participants analyzed
Hutt et al. (2017)	Detect mind wandering during a lecture using eye tracking	Eye tracker, questionnaire	Eye tracker data, questionnaire data	Yes	N/A	Bayesian network classifier	32 undergraduate students from a Canadian university
Jayaprakash, Moody, Lauría, Regan, and Baron (2014)	Detect students who are in academic difficulty	LMS, student information system	Activity log data, partial course grades, course data, student background information (× 2)	Yes	Automatic. Uses Pentaho Business Intelligence Data Integration (ETL approach)	Logistic regression, support vector machines, J48, naive Bayes	15,150 undergraduate students
Rodríguez-Triana, Prieto, Martínez-Monés, Asensio-Pérez, and Dimitriadis (2018)	Identify deviations between the desired learning state (based on learning design) and the actual state in blended/CSCL scenarios	LMS, wiki, online writing application, attendance list, human observation, instructional design information, questionnaire	Activity log data (× 2), attendance information, teacher comments, instructional design information, questionnaire data	Yes	Automatic (at least in part). Third-party tools were integrated into virtual learning environment (GLUE)	N/A (three binary classifiers were built to identify deviations between desired learning state and actual state)	165 students
Gray, McGuinness, Owende, and Hofmann (2016)	Predict at-risk students	Student information system, questionnaire, exam results	Student background information, questionnaire data, GPA	Yes	N/A	Correlations, *t* test/ANOVA. Classification (e.g., naive Bayes, decision trees)	1207 first-year students (records from 2010 to 2012)
Wang, Paquette, and Baker (2014)	Identify career path for MOOC learners	MOOC, organization member information	Student background information, questionnaire, organization member information	Yes (partly, questionnaire is analyzed separately)	N/A (most likely manual)	Chi-square, descriptive statistics	N/A (536 MOOC participants answered questionnaire)
Mangaroska, Vesin, and Giannakos (2019)	Predict student performance	E-learning portal (× 2), Integrated Development Environment (IDE)	Performance test data, activity log data (× 3)	Yes	Automatic. System collects and aggregates data from different sources. Data are integrated in a Learning Record Store	Descriptive statistics, Spearman correlation, linear regressions, visualization	21 (one teacher and 20 computer science students)

Table 3 Overview of the reviewed publications *(Continued)*

Publication	Issue addressed in publication	Types of data sources	Types of data	Data sources integrated?	Data integration approach (manual or automatic)	Methods used	Records for how many participants analyzed
Villano, Harrison, Lynch, and Chen (2018)	Examine the relationship between student retention and an early alert system (controlling for a number of variables)	Student information system, early alert system	Student background information, final grades, workload, school data (e.g., location, fee), early alert system data	Yes	Automatic. University collects and integrates data from different IT systems in a data warehouse	Survival analysis	N/A (16,142 records captured from 2011 to 2013 were analyzed)
Wong, Kwong, and Pegrum (2018)	Examine if an augmented reality app for integrity and ethics can help change student's perspectives on these subject matters	AR platform, LMS	Activity log data, text (× 2)	No (data are analyzed separately)	Data/data sources are not integrated	Descriptive statistics, text analysis, visualization	N/A (1259 students participated, but not all participants' data were included in the subsequent analyses)
Sandoval, Gonzalez, Alarcon, Pichara, and Montenegro (2018)	Prediction of students who are at risk of failing classes	Student information system, LMS	Student background information, final grades, activity log data	Yes	Automatic. Extract data from data sources and encrypt, then re-codify some of the attributes into similar types before integrating in a relational database	Linear regressions, random forest	21,314 students (over three semesters)
Sun, Xie, and Anderman (2018)	Examine the effect of self-regulation on academic achievement in flipped classrooms	Questionnaire, LMS	Questionnaire data, performance test data, partial course grades	Yes	Manual. Combine grades obtained from instructors with survey data	Structural equation modeling, multi-level regression	151 US undergraduate students
Giannakos, Sharma, Pappas, Kostakos, and Velloso (2019)	Examine if including physiological sensing data provides advantages for predicting skill acquisition (and more generally for the design of learning technologies)	Eye tracker, physiological signals wristband, EEG cap, video, game	Eye tracker data, physiological arousal data, EEG data, video derived data, activity log data, performance test data	Yes	Automatic. The features for each data source were extracted separately, then data were integrated using R	LASSO regression, random forest, ANOVA	17 participants from a major European university

Table 4 Data sources and data integration

Observation	Frequency	Publications
Multiple data sources, same format	14	Lopez Guarin et al. (2015), Park et al. (2016), Di Mitri et al. (2017), Rodríguez-Triana et al. (2018), Liu et al. (2017), Gray et al. (2016), Hutt et al. (2017), Jayaprakash et al. (2014), Raca et al. (2016), Mangaroska et al. (2019), Villano et al. (2018), Sandoval et al. (2018), Sun et al. (2018), Giannakos et al. (2019)
Multiple data sources, different formats	6	Thompson et al. (2014), Zheng et al. (2017), Pardos and Kao (2015), Ochoa et al. (2018), Wang et al. (2014), Wong et al. (2018)
Support educational data specifications	3	Pardos and Kao (2015), Di Mitri et al. (2017), Mangaroska et al. (2019)

of the data integration approach in the text. In the case where this information was not explicitly stated, we would sometimes make cautious assumptions in Table 3 (e.g., marked "N/A - most likely manual") based on the data being integrated (e.g., only integrating two lists of names could be achieved easily and more efficiently in Excel than programmatically). In one instance (Villano et al., 2018), the type of data source was not explicitly stated, in which case we had to infer it based on the types of data the system provided.

For the review, we counted one tool or system as one data source (possibly with multiple types of data). Stand-alone digital sources (e.g., corpora, list, questionnaire, and video) were also registered as data sources. Other data sources included equipment with sensors. For types of data sources and data, we iteratively defined broader categories, to allow for comparisons of data sources and data types among the studies reported in the different publications. This meant that for some of the publications, we changed data source and/or type into a broader category. For instance, the sources "Academic Information System" and "Direction of Admissions" were upon further inspection both changed into the broader category "Student Information System," and the type of data "Teamwork questionnaire" was later changed into the broader category "Questionnaire".

Table 4 presents the main observations from the reviewed literature with regard to data sources and data integration. Fourteen of the publications report on studies that combine data that are already available in the same format but come from different data sources. Six of the 20 reviewed publications report on studies that analyze data of different formats that originate from different sources, without a common format. Thus, these data are not integrated but rather analyzed separately. Only three of the studies (Di Mitri et al., 2017; Mangaroska et al., 2019; Pardos & Kao, 2015) in the reviewed publications support educational data specifications. Di Mitri et al. (2017) and Mangaroska et al. (2019) combined data with the same format, while Pardos and Kao (2015) analyzed data from multiple data sources with different formats.

Information about the data sources and types of data that occur most often in the reviewed literature are presented in Figs. 2 and 3. The most used data source is learning management system (LMS), which is used in eight of the studies reported in the reviewed publications. Student information systems and questionnaires are used as data sources in six of the studies; video is used in three; while MOOC, physiological signals wristband, and eye trackers are used in two. The majority of data sources appear only in one of the studies reported in the reviewed publications, as seen in Fig. 2.

The most used type of data is activity log, which is used in ten of the studies reported in the publications. Student background information (e.g., demographics,

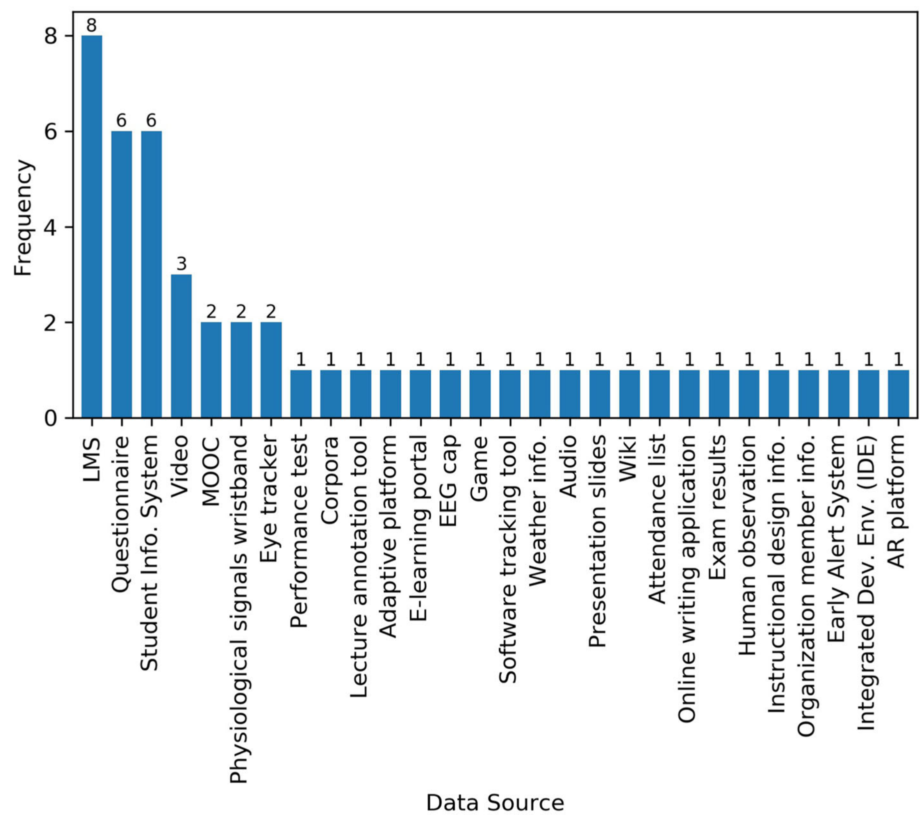

Fig. 2 Data source frequency

socio-economic information, prior academic performance) is used in eight of the studies reported in the publications, questionnaire data is used in seven, performance test data is used in five, while final grades and video derived data are used in three of the studies. Course data, text, physiological arousal data, eye tracker data, and partial course grades are each used in two of the studies reported in the publications. The majority of data types appear only in one of the reviewed publications, as seen in Fig. 3.

Table 5 shows the number of data sources used in the studies reported in each of the publications. As seen, the majority (13 of 20) use only two data sources. The other studies use three, four, five, and seven data sources, respectively.

As seen in Table 3, for data integration, 10 of the reviewed publications report on studies that use automatic integration (using tools). Manual integration requires human effort for the process of combining data, for instance, through copying and pasting data from data sources into excel sheets. Only one publication clarifies that the data are integrated manually (Sun et al., 2018). Five studies do not specify how data integration is achieved. The rest of the reported studies do not integrate data at all. Regarding issues addressed in the reviewed literature, the reported studies address a wide variety of issues. A common theme is identifying struggling students (Gray et al., 2016; Jayaprakash et al., 2014; Lopez Guarin et al., 2015; Sandoval et al., 2018).

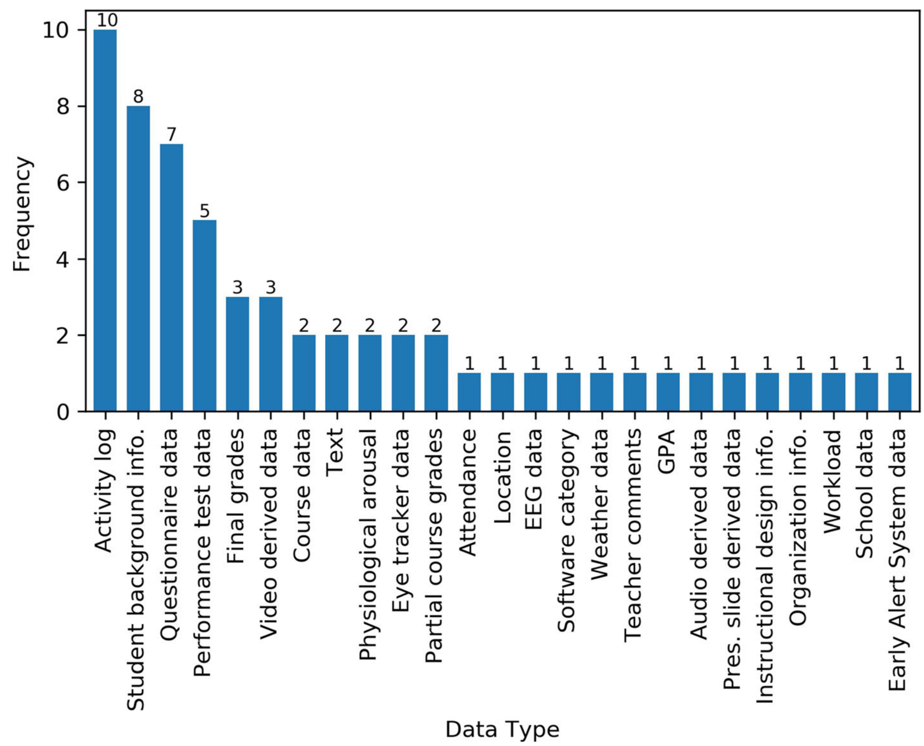

Fig. 3 Data type frequency

As shown in Table 6, different data analysis methods were used for the studies reported in the publications. The most common methods were classification (used in eight publications), regression (used in six publications), visualization (used in five publications), correlation (used in four publications), and descriptive statistics (used in three publications). Two publications (Rodríguez-Triana et al., 2018; Zheng et al., 2017) did not specify methods used in their reported studies. The number of participants in the research studies, whose data were analyzed, vary widely, ranging from 9 to 21,314 (see Table 3). Six of the publications (Pardos & Kao, 2015; Park et al., 2016; Villano et al., 2018; Wang et al., 2014; Wong et al., 2018; Zheng et al., 2017) did not list the number of participants whose data were analyzed in their reported studies.

Table 5 Number of data sources

Number of data sources	References
2	Lopez Guarin et al. (2015), Park et al. (2016), Thompson et al. (2014), Zheng et al. (2017), Pardos and Kao (2015), Raca et al. (2016), Hutt et al. (2017), Jayaprakash et al. (2014), Wang et al. (2014), Villano et al. (2018), Wong et al. (2018), Sandoval et al. (2018), Sun et al. (2018)
3	Ochoa et al. (2018), Gray et al. (2016), Mangaroska et al. (2019)
4	Liu et al. (2017), Di Mitri et al. (2017)
5	Giannakos et al. (2019)
7	Rodríguez-Triana et al. (2018)

Table 6 Methods used

Method used	Publications
Classification methods	Lopez Guarin et al. (2015), Park et al. (2016), Ochoa et al. (2018), Hutt et al. (2017), Jayaprakash et al. (2014), Gray et al. (2016), Sandoval et al. (2018), Giannakos et al. (2019)
Regression methods	Liu et al. (2017), Di Mitri et al. (2017), Mangaroska et al. (2019), Sandoval et al. (2018), Sun et al. (2018), Giannakos et al. (2019)
Visualization	Thompson et al. (2014), Pardos and Kao (2015), Liu et al. (2017), Mangaroska et al. (2019), Wong et al. (2018)
Correlation	Liu et al. (2017), Raca et al. (2016), Gray et al. (2016), Mangaroska et al. (2019)
Descriptive statistics	Raca et al. (2016), Mangaroska et al. (2019), Wong et al. (2018)
ANOVA	Gray et al. (2016), Giannakos et al. (2019)
Text analysis	Thompson et al. (2014), Wong et al. (2018)
t test	Gray et al. (2016)
Network analysis	Pardos and Kao (2015)
Chi Square	Wang et al. (2014)
Structural equation modelling	Sun et al. (2018)
Survival analysis	Villano et al. (2018)

There is also some variation (e.g., relational database, text file, learning record store) in the reported data storage technologies used, as seen in Table 7. However, out of the 20 publications, only seven report the type of technology used in the studies.

Results summary and discussion

Having analyzed the reviewed literature, we can now address the three research questions proposed in this paper. A number of other patterns and trends have emerged in relation to LA in terms of data integration in the context of higher education; these patterns and trends are also discussed in the following.

Re-visiting research questions

What types of data and data sources are being used for LA in higher education?

In the reviewed literature, we see that three data sources appear in many of the research studies, namely LMSs, student information systems, and questionnaires. The use of LMSs and student information systems is most likely because of the wide use of these systems in higher education institutions and the amount of digital information that can be extracted from them. In the case where data are not already collected through systems, questionnaires allow stakeholders to self-report the information. The most common types of data identified are activity logs, student background information, questionnaire data, and

Table 7 Data storage technologies

Data storage technology	Publication
Relational database	Lopez Guarin et al. (2015), Sandoval et al. (2018)
Text file	Thompson et al. (2014)
Comma separated values file	Pardos and Kao (2015)
Learning record store	Di Mitri et al. (2017), Mangaroska et al. (2019)
Data warehouse	Villano et al. (2018)

performance test data. This is perhaps not surprising given that these types of data are often collected in the data sources that have been found to be the most common.

From 2017, there seems to be a shift, where multimodal data are used to a larger extent in the studies (Di Mitri et al., 2017; Giannakos et al., 2019; Ochoa et al., 2018; Rodríguez-Triana et al., 2018). Diverse data are collected from data sources containing sensors, such as physiological signals wristbands, eye trackers, and EEG caps. The use of sensor data in the research, however, is generally still at an early stage, facing challenges such as synchronization in addition to data integration.

How and to what extent are different data being used/combined for LA research in higher education?

The results indicate that the majority of the studies in the reviewed publications that combine data from different sources use tools for automatic integration. The tools used include Business Intelligence software (Jayaprakash et al., 2014), tools developed ad hoc for a specific research project (Di Mitri et al., 2017), SQL (Lopez Guarin et al., 2015), and R scripts (Giannakos et al., 2019). However, out of sixteen studies reported in the publications that integrate data, five of the studies do not specify the integration approach, while one specifies that the approach is manual. Most publications report on studies that combine data of similar formats. A number of the studies that use multiple data sources do not integrate data but rather analyze them separately.

The results also show that the number of data sources used tends to be limited. The majority of studies reported in the publications use two data sources. Only seven out of 20 used more than two data sources. As mentioned earlier, we also see that a small number of data sources appear in many of the research studies. However, higher education often includes a broad variety of tools, such as digital exam systems, library systems, and key card access systems. When only a limited number of data sources are taken into account for LA, this may give only a partial picture of student behavior and learning, thereby also potentially biasing analysis results. On the other hand, when increasing the number of data sources, the complexity of the problem of interoperability is also increased. Challenges such as meaningful data integration, storage, and processing requirements need to be addressed.

What methods are used, and what issues are being addressed through using/combining multiple data sources?

As the results have shown, the most common methods used are classification methods. This is in accordance with findings in previous research (Shahiri et al., 2015). Regression methods, visualization, correlation, and descriptive statistics are also used to analyze data in a number of the publications. While the issues addressed in the reviewed literature vary greatly, a common theme is identifying struggling students. The publications addressing this theme of issues have been published between 2014 and 2018, indicating that addressing this type of theme has been of interests for a few years. It is clear that identifying and helping students who are at risk can have positive consequences, not only for the individual students, but also for the higher education institutions (e.g., in terms of economy via student throughput and reputation).

Additional findings

Lack of details in technical solutions

The reviewed literature tends to lack details about the technical solutions for combining multiple data sources in the implemented systems. There is generally little information about data formats, data storage solutions, and data integration techniques. Only four publications report on studies that integrate data sources and provide information (implicitly or explicitly) on all three of formats, storage solutions, and integration techniques. Lopez Guarin et al. (2015) detail a data integration technique of joining database tables, which implies the use of a relational database for storage, with the relational model as the underlying format. Sandoval et al. (2018) also integrate data into a relational database. Di Mitri et al. (2017) and Mangaroska et al. (2019) format data using the xAPI specification. The collected data are stored and integrated in a Learning Record Store. For the five publications that do not specify the data integration technique used in the studies, we may speculate that they are combining the data manually. However, without any specific information, it is hard to draw any hard conclusions.

Lack of involvement of stakeholders

Most of the reviewed publications have no information regarding who chose, combined, and managed the data used in the studies. This indicates a lack of involvement of stakeholders and adoption of participatory or co-design approaches. Successfully involving and engaging different stakeholders may be a challenge at the institutional level, both in terms of management and organization (Buckingham Shum & McKay, 2018; Dawson et al., 2018). The reviewed publications tend to make clear that it was the researchers themselves who did the feature selection and analysis, while further data management is not addressed. One exception is found in the publication by Rodríguez-Triana et al. (2018). Here, a teacher collaborated with the researchers to customize a multimodal LA solution for blended learning, with emphasis on the data-gathering phase (e.g., what questions are to be answered by the solution and what data are to be used given constraints and affordances).

Lack of use of educational data specifications

Using educational data specifications such as xAPI and IMS Caliper Analytics enables a more uniform representation of activity data, thus supporting the combination of data that follow the same standards. In the reviewed publications, however, we only find evidence of three studies using the xAPI specification (Di Mitri et al., 2017; Mangaroska et al., 2019; Pardos & Kao, 2015), while none use Caliper Analytics. Even though xAPI and Caliper are mainly focused on the technical level of interoperability (with added capabilities for the semantic level), there are gaps in the specifications that can make it challenging to describe parts of learning and learning environments, and inconsistent use of a specification among different communities of practice may make it difficult to efficiently integrate data, these specifications are still a good starting point for data integration in LA. It is not clear why xAPI and Caliper Analytics seem to be used so seldom in LA research. One possible answer may be that those implementing LA solutions lack skills in using such specifications. With regards to some individual data sources, e.g., digital tools such as LMSs, it may also be a factor that EdTech/digital tool vendors do not tend to provide data in xAPI or Caliper Analytics formats, in which case, those using educational data specifications need to transform data from the original format to the relevant standardized format themselves. It is clear that this transformation effort can be time-consuming, even for LA solution implementers that are skilled using the educational data specifications (for some digital tools, there are plugins that may help with the transformation process [JISC, 2019]). In addition, it may also be a factor that some of the usage of xAPI and Caliper has not been illustrated through this review, as there are few publications included on platforms. Further examination is needed in order to understand why educational data specifications do not seem to be widely used for data integration in LA.

Combination of similar data formats as low-hanging fruit

In the reviewed literature, we see that most publications report on studies that combine data of similar formats, as mentioned earlier. For some of these studies, the data did not originally have a common format; for instance, Raca et al. (2016) get data from the diverse sources of video and questionnaires. Other studies combine data that are already in similar formats. For instance, Di Mitri et al. (2017) and Rodríguez-Triana et al. (2018) both get data from application programming interfaces (APIs), which tend to output JSON format. The transformation effort is trivial when data are originally in the same format, but more challenging when the formats are not originally equivalent. Thus, focusing on data already in the same format is a low-hanging fruit in terms of data integration. However, there is still the significant challenge of semantics, ensuring the data have the same meaning. In general, the alignment of concepts is usually done ad hoc when programming an analytics solution. Another approach to the alignment of concepts would be to use an ontology, "a formal, explicit specification of a shared conceptualization" (Studer, Benjamins, & Fensel, 1998, p. 184). Ontologies are related to controlled vocabularies, although in general, they tend to have more complex and formal collections of concepts and relationships (W3C, 2015). When using an ontology, it is possible to add descriptions and meaning to data coming from various sources, and to combine, support, and reuse different specifications (Allemang & Hendler, 2011). Ontologies also enable inference, meaning we can state new and related facts from one stated fact.

Recommendations and conclusion

This review has examined LA research in higher education that uses and/or combines data from multiple sources, aiming to assess the current state of LA in terms of data integration in the context of higher education.

We acknowledge that there are some limitations in this systematic literature review. Limiting the language to English may have excluded relevant research published in other languages. Choosing to focus on LA in the context of higher education has excluded some research that use multiple data sources which take place in other levels of education, such as K-12 (Chang et al., 2017; Mutahi et al., 2015). While the research studies were excluded for not taking place in higher education, both examine how students engage with interactive learning content, whether individually or in groups. The information provided and research focus varies for the publications included in the review. Some publications lack information that we consider important for the review (lacking information is marked N/A in Table 3). While we are aware that there are LA platforms such as OnTask (2019), JISC (2019), and Apereo (2018) that address data integration issues, the review process only resulted in two publications with research where it is clear that they were using platforms (Mangaroska et al., 2019; Pardos & Kao, 2015), leading us to believe that the published research available on such solutions is limited. It is possible, however, that some of the research studies were using platforms without stating it clearly, since many of the publications did not provide much detail about the technical solutions.

While the choice of data sources used in research depends of the purpose of the learning analytics (e.g., identifying dropouts or successful students, giving feedback), we reason that the actual choice of sources might also follow from access to data sources not always being that easy (this is the case from our own experience as well). In addition, integrating data from multiple sources is a challenge. Thus, based on the finding that few studies actually use and combine multiple data sources, we identify a need for more research studies that use and combine more data sources and report the details so others can learn from the experience (including how they managed to access and acquire the data).

Integrating multiple data sources plays an important role in scaling up LA. The problem of scalability in terms of LA in higher education can be viewed as a natural extension of the problem of data integration. In addition to the problem of combining data from multiple sources, it encompasses additional dimensions such as the institutional (the solution should at minimum scale across departments and faculties within a higher education institution) and the temporal (the solution should scale across semesters and years, to allow for a more complete picture of learning). In higher education, there are multiple tools that generate data, and being able to integrate these data can promote the usefulness of LA to a different level than one data source can do by itself. If the studies conducted are to result in realistic analysis results, it is important that they reflect learning and learning environments as realistically as possible; thus, it may be beneficial to combine more than one or a limited set of data sources.

With regard to the problem that the reported studies use limited data sources, we argue that institutional policy should guide data use. To break down data silos and enable more efficient application of diverse data sources, such policy should cover not only the technical levels of data exchange, but also factors related to organization and management.

Privacy, which is guided by national and international laws, is essential for LA in general and for data integration in particular. Securing user data includes prevention of unauthorized access to the systems. In addition, de-identification of user information yields added privacy. When aggregating de-identified data from multiple sources, there is the added chance that users can become indirectly identifiable from the collective information that is stored. On the other hand, there is the risk that ensuring data privacy may limit access to data that could be important for LA research (Flanagan & Ogata, 2017). Such considerations need to be balanced, complying with relevant laws.

Having conducted the review, we found that in general researchers do not describe their process of data integration in enough detail (or at all). Our recommendation is that they make this important part of LA research clearer, as building on existing knowledge can help push the boundaries of the state of the art and help in replication of earlier studies.

As mentioned in the discussion, it is not clear why xAPI and Caliper Analytics seem to be so seldom used, even though these specifications help enable interoperability on the technical level and have capabilities for the semantic level. Here, we would suggest that studies look further into the specifications and their usage, in order to understand the challenges in adoption and the opportunities afforded, from the LA stakeholders' point of view.

In terms of stakeholder participation, we would recommend more inclusion of different stakeholders in LA research. To really understand the pressing issues, such as how to interpret analysis results in a given context, it is not enough to include only researchers and technical staff. Those that have more thorough knowledge of the problem domain (e.g., teachers for classroom studies or instructors for university course studies) also need to be included. There also needs to be more focus on participatory design or co-design, where stakeholders with diverse competences and strengths use their individual knowledge for implementation and scalability of LA in higher education. Finally, there needs to be a dialog between the LA research community and the EdTech/digital tool vendor sector about the need to have standardized data.

Abbreviations
API: Application Programming Interface; EDM: Educational Data Mining; LA: Learning Analytics; LMS: Learning Management System; MOOC: Massive Open Online Course

Acknowledgements
There are no acknowledgements.

Authors' contributions
The systematic review presented in this article is a part of the PhD project conducted by JS. BW is her main supervisor, and WC is her co-supervisor. WC has been working closely with JS on judging the relevance of each identified publication for the review, and assisted in planning the review and article writing. BW has contributed to judging the relevance of problematic publications, planning the review, and article writing. All authors read and approved the final manuscript.

Funding
This research is part of JS's PhD funded by the Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway.

Availability of data and materials
Data for the review was extracted from publications identified through searching academic databases and a limited number of conference proceedings/journals for publishing LA research.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Centre for the Science of Learning & Technology, University of Bergen, Bergen, Norway. [2]Department of Information Science & Media Studies, University of Bergen, Bergen, Norway. [3]Oslo Metropolitan University, Oslo, Norway.

Received: 25 January 2019 Accepted: 26 July 2019
Published online: 28 August 2019

References

Allemang, D., & Hendler, J. (2011). *Semantic web for the working ontologist: effective modeling in RDFS and OWL* (2nd ed.). Morgan Kaufmann.

Apereo. (2018). Learning Analytics Initiative | Apereo. Retrieved from https://www.apereo.org/communities/learning-analytics-initiative

Buckingham Shum, S., & McKay, T. (2018). Architecting for learning analytics: Innovating for sustainable impact. *EDUCAUSE Review*.

Chang, C.-J., Chang, M.-H., Liu, C.-C., Chiu, B.-C., Fan Chiang, S.-H., Wen, C.-T., et al. (2017). An analysis of collaborative problem-solving activities mediated by individual-based and collaborative computer simulations: Collaborative problem solving. *Journal of Computer Assisted Learning, 33*(6), 649–662 https://doi.org/10.1111/jcal.12208.

Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an open learning analytics ecosystem. In B. K. Daniel (Ed.), *Big Data and learning analytics in higher education* (pp. 195–219) https://doi.org/10.1007/978-3-319-06520-5_12.

Cooper, A., & Hoel, T. (2015). *Data sharing requirements and roadmap*.

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). *Rethinking learning analytics adoption through complexity leadership theory* (pp. 236–244). ACM Press https://doi.org/10.1145/3170358.3170375.

Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). *Learning pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data* (pp. 188–197). ACM Press https://doi.org/10.1145/3027385.3027447.

European Commission. (2017). New european interoperability framework. Retrieved from https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf

European Union. (2016). Regulations. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Flanagan, B., & Ogata, H. (2017). Integration of learning analytics research and production systems while protecting privacy. In *The 25th International Conference on Computers in Education, Christchurch, New Zealand* (pp. 333–338).

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management, 48*, 108–119 https://doi.org/10.1016/j.ijinfomgt.2019.02.003.

Gray, G., McGuinness, C., Owende, P., & Hofmann, M. (2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics, 3*(2), 330–372 https://doi.org/10.18608/jla.2016.32.20.

Hutt, S., Hardey, J., Bixler, R., Stewart, A., Risko, E., & D'Mello, S. K. (2017). Gaze-based detection of mind wandering during lecture viewing. In *Proceedings of the 10th International Conference on Educational Data Mining* (pp. 226–231).

IMS Caliper Analytics. (2019). Caliper Analytics | IMS Global Learning Consortium. Retrieved from https://www.imsglobal.org/activity/caliper

Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: an open source analytics initiative. *Journal of Learning Analytics, 1*(1), 6–47 https://doi.org/10.18608/jla.2014.11.3.

JISC. (2019). Learning records warehouse: technical overview: Integration overview. Retrieved from https://docs.analytics.alpha.jisc.ac.uk/docs/learning-records-warehouse/Technical-Overview:%2D%2DIntegration-Overview

Kadadi, A., Agrawal, R., Nyamful, C., & Atiq, R. (2014). Challenges of data integration and interoperability in big data. *2014 IEEE International Conference on Big Data (Big Data)*, 38–40 https://doi.org/10.1109/BigData.2014.7004486.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, Keele University (UK).

Liu, M., Kang, J., Zou, W., Lee, H., Pan, Z., & Corliss, S. (2017). Using data to understand how to better design adaptive learning. *Technology, Knowledge and Learning, 22*(3), 271–298 https://doi.org/10.1007/s10758-017-9326-z.

Lopez Guarin, C. E., Guzman, E. L., & Gonzalez, F. A. (2015). A model to predict low academic performance at a specific enrollment using data mining. *IEEE Revista Iberoamericana de Tecnologias Del Aprendizaje, 10*(3), 119–125 https://doi.org/10.1109/RITA.2015.2452632.

Mangaroska, K., Vesin, B., & Giannakos, M. (2019). Cross-platform analytics: A step towards personalization and adaptation in education. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge - LAK19* (pp. 71–75) https://doi.org/10.1145/3303772.3303825.

Misiejuk, K., & Wasson, B. (2017). *State of the field report on learning analytics*. Bergen: *SLATE Report 2017-2* Retrieved from http://bora.uib.no/bitstream/handle/1956/17740/SoF%20Learning%20Analytics%20Report.pdf.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA group. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement* (p. 7).

Mutahi, J., Bent, O., Kinai, A., Weldemariam, K., Sengupta, B., & Contractor, D. (2015). Seamless blended learning using the cognitive learning companion: A systemic view. *IBM Journal of Research and Development, 59*(6), 8:1–8:13 https://doi.org/10.1147/JRD.2015.2463591.

Ochoa, X., Domínguez, F., Guamán, B., Maya, R., Falcones, G., & Castells, J. (2018). *The RAP system: automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors* (pp. 360–364). ACM Press https://doi.org/10.1145/3170358.3170406.

OnTask. (2019). Home | OnTask. Retrieved from https://www.ontasklearning.org/

Pardos, Z. A., & Kao, K. (2015). *moocRP: An open-source analytics platform* (pp. 103–110). ACM Press https://doi.org/10.1145/2724660.2724683.

Park, Y., Yu, J. H., & Jo, I.-H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education, 29*, 1–11 https://doi.org/10.1016/j.iheduc.2015.11.001.

Raca, M., Tormey, R., & Dillenbourg, P. (2016). Sleepers' lag: Study on motion and attention. *Journal of Learning Analytics, 3*(2), 239–260 https://doi.org/10.18608/jla.2016.32.12.

Rodríguez-Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2018). *The teacher in the loop: Customizing multimodal learning analytics for blended learning* (pp. 417–426). ACM Press https://doi.org/10.1145/3170358.3170364.

Sandoval, A., Gonzalez, C., Alarcon, R., Pichara, K., & Montenegro, M. (2018). Centralized student performance prediction in large courses based on low-cost variables in an institutional context. *The Internet and Higher Education, 37*, 76–89 https://doi.org/10.1016/j.iheduc.2018.02.002.

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education: A review of UK and international practice. Jisc. Retrieved from https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v3.pdf

Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science, 72*, 414–422 https://doi.org/10.1016/j.procs.2015.12.157.

Siemens, G. (2011). 1st international conference on learning analytics and knowledge. Technology Enhanced Knowledge Research Institute (TEKRI). Retrieved from https://tekri.athabascau.ca/analytics/

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *Data and knowledge engineering, 25*(1), 161–198.

Sun, Z., Xie, K., & Anderman, L. H. (2018). The role of self-regulated learning in students' success in flipped undergraduate math courses. *The Internet and Higher Education, 36*, 41–53 https://doi.org/10.1016/j.iheduc.2017.09.003.

Thompson, K., Kennedy-Clark, S., Wheeler, P., & Kelly, N. (2014). Discovering indicators of successful collaboration using tense: Automated extraction of patterns in discourse: Discovering indicators of successful collaboration. *British Journal of Educational Technology, 45*(3), 461–470 https://doi.org/10.1111/bjet.12151.

Villano, R., Harrison, S., Lynch, G., & Chen, G. (2018). Linking early alert systems and student retention: a survival analysis approach. *Higher Education, 76*(5), 903–920 https://doi.org/10.1007/s10734-018-0249-y.

W3C. (2015). Ontologies. Retrieved from https://www.w3.org/standards/semanticweb/ontology

Wang, Y., Paquette, L., & Baker, R. (2014). A Longitudinal Study on Learner Career Advancement in MOOCs. *Journal of Learning Analytics, 1*(3), 203–206 https://doi.org/10.18608/jla.2014.13.23.

Wong, E. Y. W., Kwong, T., & Pegrum, M. (2018). Learning on mobile augmented reality trails of integrity and ethics. *Research and Practice in Technology Enhanced Learning, 13*(1) https://doi.org/10.1186/s41039-018-0088-6.
xAPI (2019). xAPI.com Homepage: What is xAPI (the Experience API). Retrieved from https://xapi.com/
Zheng, M., Bender, D., & Nadershahi, N. (2017). Faculty professional development in emergent pedagogies for instructional innovation in dental education. *European Journal of Dental Education, 21*(2), 67–78 https://doi.org/10.1111/eje.12180.

Publisher's Note

Paper 2

Samuelsen, J., Chen, W., & Wasson, B. (2021). Enriching context descriptions for enhanced LA scalability: a case study. *Research and Practice in Technology Enhanced Learning*, *16*(1). https://doi.org/10.1186/s41039-021-00150-2

Research and Practice in Technology Enhanced Learning

RESEARCH Open Access

Enriching context descriptions for enhanced LA scalability: a case study

Jeanette Samuelsen[1,2*], Weiqin Chen[1,3] and Barbara Wasson[1,2]

* Correspondence: Jeanette.Samuelsen@uib.no
[1]Centre for the Science of Learning & Technology, University of Bergen, P.O. Box 7807, 5020 Bergen, Norway
[2]Department of Information Science & Media Studies, University of Bergen, P.O. Box 7802, 5020 Bergen, Norway
Full list of author information is available at the end of the article

Abstract

Learning analytics (LA) is a field that examines data about learners and their context, for understanding and optimizing learning and the environments in which it occurs. Integration of multiple data sources, an important dimension of scalability, has the potential to provide rich insights within LA. Using a common standard such as the Experience API (xAPI) to describe learning activity data across multiple sources can alleviate obstacles for data integration. Despite their potential, however, research indicates that standards are seldom used for integration of multiple sources in LA. Our research aims to understand and address the challenges of using current learning activity data standards for describing learning context with regard to interoperability and data integration. In this paper, we present the results of an exploratory case study involving in-depth interviews with stakeholders having used xAPI in a real-world project. Based on the subsequent thematic analysis of interviews, and examination of xAPI, we identified challenges and limitations in describing learning context data, and developed recommendations (provided in this paper in summarized form) for enriching context descriptions and enhancing the expressibility of xAPI. By situating the research in a real-world setting, our research also contributes to bridge the gap between the academic community and practitioners in learning activity data standards and scalability, focusing on description of learning context.

Keywords: Learning context, Learning activity data specification, xAPI, Scalability, Interoperability, Data integration, Learning analytics

Introduction

A diversity of digital tools are used within education. They may, for instance, facilitate exam taking (e.g., an exam system), store student demographic and result data (e.g., a student information system), or make available lecture notes and videos (e.g., a learning management system [LMS]). When a student uses such systems, digital trace data may be generated and saved in individual data sources. This data can be used to gain insight into the student and their learning. Learning analytics (LA) is "the 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" (Siemens, 2011). Within LA, "researchers have suggested that the true potential to offer meaningful insight comes from combining data from across different sources" (Bakharia, Kitto, Pardo, Gašević, & Dawson, 2016, p. 378). Data integration, the combination of data from different sources, also plays an important role for the scalability of LA (Samuelsen, Chen, & Wasson, 2019).

Throughout their various activities, learners are situated in different contexts. They move within the physical space, at varying times of the day, using different tools on separate devices, leading to data being generated through different sensors. A context is defined as "any information that can be used to characterize the situation of an entity" (Dey, 2001, p. 5), and an entity can be a person, place, or an object (Dey, 2001). To integrate data from different sources in LA, it is crucial to take into account the context of the data. Taking context into account can have benefits for interoperability and may also be used to personalize learning for the individual learner, as well as enable better querying and reporting of the data.

Data integration is closely related to interoperability, which involves semantic, technical, legal, and organizational levels (European Commission, 2017). Concerning technical and semantic interoperability, two well-known data specifications (de facto industry standards) exist which target the educational domain and LA, namely the Experience API (xAPI; Advanced Distributed Learning, 2017a) and IMS Caliper Analytics (2020). These specifications both enable the exchange and the integration of learning activity data originating from different tools and data sources, where individual activity data describe a learner interacting with a learning object in a learning environment (modelled with the most basic structure of "actor verb object"). The activity data can subsequently be stored in a Learning Record Store (LRS). Both specifications also provide mechanisms for adding vocabularies, through profiles, which can help in terms of structuring the activity data and adding semantics. Current profile specifications are specified in the JSON for Linking Data (JSON-LD)[1] format, which builds on JSON and semantic technologies to enable machine-readable data definitions. For xAPI, any community of practice can create a new profile, while for Caliper only organizations that are members of IMS may contribute to profiles (and other parts of the specification). As the latter may suggest, the usage of xAPI is generally more flexible than that of Caliper (Griffiths & Hoel, 2016). For a detailed comparison of xAPI and Caliper, please refer to Griffiths and Hoel (2016).

Despite the availability of these learning activity data specifications, previous research (Samuelsen et al., 2019) found that they are not widely used for data integration of data coming from multiple data sources for LA in the context of higher education; in the case of xAPI, a few examples of use were found, while no examples of Caliper use were found. Thus, it should be of interest for researchers and practitioners in LA to know why there seems to be so little use of learning activity data specifications, and to understand the challenges and limitations when using the existing specifications.

This paper reports on an exploratory case study where we look at the challenges and limitations of using a current learning activity data standard (i.e., xAPI) for describing the learning context. While previous research has identified some of the challenges and limitations of xAPI (Bakharia et al., 2016; Betts & Smith, 2019; Keehn & Claggett, 2019), to our knowledge no studies have systematically collected and analyzed data from stakeholders who have used xAPI in a real-world case and identified the gaps

[1] https://json-ld.org/

between xAPI and the needs of stakeholders with regard to interoperability and integration of learning activity data. Thus, we aim to contribute to the knowledge base through the systematic collection, analysis and identification of xAPI gaps and needs with regard to interoperability and data integration as they have been experienced by stakeholders in a real-world case.

The case is the Activity data for Assessment and Adaptation (AVT) project (Morlandstø, Hansen, Wasson, & Bull, 2019), a Norwegian project exploring the use of activity data coming from multiple sources to adapt learning to individual learner needs and for use in assessment. AVT used the xAPI data specification for describing student activity data originating from different sources. Through in-depth interviews with AVT stakeholders with varying perspectives, and inspection of the xAPI specification (Advanced Distributed Learning, 2017a) and the xAPI profiles specification (Advanced Distributed Learning, 2018a), we identified some challenges and limitations of xAPI, focusing on learning context description. Based on the identified challenges and limitations, we have provided recommendations on how xAPI can be improved to enhance its expressibility—meaning it should be possible to describe data in a consistent way across data sources—in order to better support interoperability, data integration and (consequently) scalability. This paper answers the following research questions:

RQ1: Focusing on descriptions of xAPI context, what are the gaps and needs regarding interoperability and data integration?
RQ2: How should the identified gaps and needs be addressed in order to provide improved interoperability and data integration?

Background

In this section, we first look at several data models that attempt to formalize context. Next, we examine the constructs currently available in xAPI that enable the description of context. Then we conclude with a comparison of the context data models and xAPI.

Context Data Models

Jovanović, Gašević, Knight, and Richards (2007) developed an ontology-based framework, Learning Object Context Ontologies (LOCO), to formalize and record context related to learning objects (i.e., digital artifacts used for learning on digital platforms). Learning objects consist of learning content and are assigned to learning activities to achieve learning objectives. The LOCO framework integrates several ontologies, e.g., for learning object content structure and user modelling. The learning object context that can be recorded, i.e., metadata which originates from a learning process, includes information about the learning object domain, the learning situation, and the learner. Two tools were developed based on the LOCO framework. The first tool, LOCO-Analyst, can generate feedback for instructors based on analysis of context data collected from an online learning environment (e.g., LMS). The second tool, TANGRAM, targets the learners and is a “Web-based application for personalized learning in the area of Intelligent information systems” (Jovanović et al., 2007, p. 57). It personalizes the assembly of learning content.

Schmitz, Wolpers, Kirschenmann, and Niemann (2011) detail a framework for collecting and analyzing contextual attention metadata (CAM) from digital environments. CAM expresses data selection behaviors of users. The authors developed a schema to represent CAM that allows for the registration of aspects such as which data objects (e.g., file, video, email message) capture the attention of users, what actions are performed on the objects (e.g., a file was opened), and what was the context of use when a user interacted with an object (e.g., time, location). To enable the collection of CAM records, the approach is to add file system/application wrappers, thereby transforming the original data format to the format of CAM in XML. The CAM schema is semi-structured for some properties, such as context. The context property is a container for data of varying types, i.e., an arbitrary number of key-value pairs can be stored within this container. The authors note that while the semi-structured approach is flexible and allows for registering different types of data, it also creates challenges for exchanging data because data can be described in different ways. They state that an alternative would be to import different metadata schema, which could be used to structure the different types of data. Such an approach would rely on pre-defined schemas, e.g., from FOAF[2] and Dublin Core[3]. To avoid redundancy of stored metadata, the authors describe a tentative approach where metadata are stored as triple representations (subject, predicate, object), and where pointers are added to other metadata descriptions. Regarding CAM, one prototype was implemented to collect, analyze and visualize user communication. Different metadata were extracted and transformed into CAM format, providing the basis for visualizing the social network of the user, including the type of communication that took place and the user's communication behavior. Another prototype using CAM was developed for an online learning environment. Here, data object metadata (e.g., number of object uses) were utilized for user recommendations. Usage and behavior data were also visualized for the individual user, adding the potential for providing metacognitive support.

The learning context project (Thüs et al., 2012) recognizes that devices, such as mobile phone and tablets that contain a diverse set of sensors, have possibilities for recording context. The project has developed the Learning Context Data Model (LCDM), a suggested standard to represent context data and enable increased interoperability and reusability of context models. The data model considers learners and events, where an event is categorized at either a higher or lower level. Available high-level categories are activity (e.g., writing a paper), environmental (e.g., noise level, location), and biological (e.g., heart rate level, level of motivation) (Muslim, Chatti, Mahapatra, & Schroeder, 2016). There are a limited number of low-level categories, and for each, the data model specifies required and recommended inputs. Context can be broadly categorized as extrinsic or intrinsic. Extrinsic context is related to the user’s current environment, while intrinsic context is related to the inside of the user such as knowledge, concentration, or motivational level (Thüs et al., 2012). The LCDM allows for the registration of both extrinsic and intrinsic context events. It can also register user interests and platforms, which specifies where an event was captured, e.g., on a mobile phone. In addition to the data model, the learning context project provides an API that enables storage and

[2]http://xmlns.com/foaf/spec/
[3]https://dublincore.org/specifications/dublin-core/

retrieval of the context-related information and visualizations for the collected data. For instance, one visualization shows how learner interests evolve over time, something that may enable self-reflection (Thüs, Chatti, Brandt, & Schroeder, 2015).

Lincke (2020) describes an approach for context modelling in her PhD dissertation, where she has developed a rich context model (RCM). The RCM approach models the user context according to specific context dimensions, each relating to a given application domain. The RCM was designed for generalizability in terms of application domains; thus, it can be utilized for different domains through providing separate configurations for the individual domains, removing the need to change the core of the model to add new domains. The configurations can specify aspects such as expected data/data types and database configuration. In the research, context was modelled for different application domains, such as mobile learning, LA, and recommender systems. For instance, in the mobile learning application domain, dimensions were modelled for environment, device, and personal context. The environment context could include contextual information such as location, weather conditions, and nearby places; the device context could include information such as screen size, battery level, and Internet connectivity; and the personal context could include information such as demographics, courses, interests, and preferences. In the dissertation by Lincke (2020), much emphasis is placed on data analysis, especially with regard to user recommendations of relevant items as they pertain to the current situation, thus offering personalization/contextualization to the user. Analysis of context data, with resulting recommendations, has been implemented in tools within the mobile learning application domain. Data analysis results were visualized in mobile learning and several other application domains.

Context in xAPI

xAPI statements are made up of the most basic building blocks of "actor verb object" (see Fig. 1 for more information on available properties and structures). An xAPI activity is a type of object that an actor has interacted with. Used together with a verb, the activity may represent a unit of instruction, performance, or experience. The interpretation of an activity is broad, meaning this concept can not only be used to represent virtual objects, but also physical/tangible objects (Advanced Distributed Learning, 2017b).

The xAPI specification (currently in version 1.0.3), expressed in the JSON format, is quite flexible. For instance, users are free to define new verbs and activity types (an activity is an instance of an activity type) for use in statements, ideally publishing these vocabulary concepts in profiles shared with relevant communities of practice. Additionally, a number of the expected value formats have a flexible structure (e.g., JSON objects may contain an arbitrary number of properties of varying levels of nesting). Finally, a number of structures/properties that can be used for data description are optional. In xAPI statements, the context structure is an optional structure that allows us to register context data. Since xAPI is a standard for learning activity data, context is related to the learner as they interact with a learning object in a (typically) digital environment. The context structure is on the same level in a statement as the actor, verb, and object structures. Another structure on this level, which also allows for registration

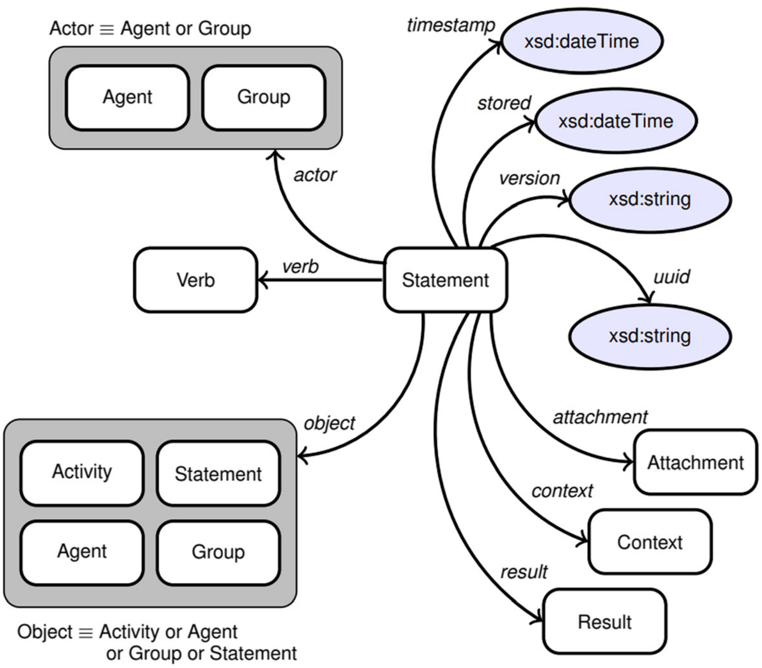

Fig. 1 xAPI statement (Vidal, Rabelo, & Lama, 2015)

of context data related to learning activity data, is the result structure that "represents a measured outcome related to the Statement in which it is included" (Advanced Distributed Learning, 2017c); it may contain information on score, duration, response, success, completion, and other relevant (user-defined) attributes. There are nine properties that can be used within the xAPI context structure (see Table 1). Seven of them are defined with keys that require a single value or object that represents a single entity, including registration (value format is a UUID), instructor (value is an agent, stored in a JSON

Table 1 Context structure properties (Advanced Distributed Learning, 2017c)

Property	Type	Description	Required
registration	UUID	The registration that the Statement is associated with.	Optional
instructor	Agent (MAY be a Group)	Instructor that the Statement relates to, if not included as the Actor of the Statement.	Optional
team	Group	Team that this Statement relates to, if not included as the Actor of the Statement.	Optional
contextActivities	contextActivities Object	A map of the types of learning activity context that this Statement is related to. Valid context types are: "parent", "grouping", "category" and "other".	Optional
revision	String	Revision of the learning activity associated with this Statement. Format is free.	Optional
platform	String	Platform used in the experience of this learning activity.	Optional
language	String (as defined in RFC 5646)	Code representing the language in which the experience being recorded in this Statement (mainly) occurred in, if applicable and known.	Optional
statement	Statement Reference	Another Statement to be considered as context for this Statement.	Optional
extensions	Object	A map of any other domain-specific context relevant to this Statement. For example, in a flight simulator altitude, airspeed, wind, attitude, GPS coordinates might all be relevant.	Optional

object), statement (value is another xAPI statement that is found to be relevant to the xAPI statement, stored in a JSON object), and team (value is an xAPI group, stored in a JSON object). The properties language, platform, and revision (of a learning activity) all require a string as their value (Advanced Distributed Learning, 2017c). Context information not suitable for these seven properties that all take a single value or object that represents a single entity, and not related to a measured outcome (i.e., result), can be described with ContextActivities and extensions.

ContextActivities let us specify "a map of the types of learning activity context that this Statement is related to" (Advanced Distributed Learning, 2017c). The available context types are *parent*, *grouping*, *category*, and *other*. The *parent* structure is used to specify the parent(s) of the object activity of a statement. For instance, a quiz would be the parent if the object of a statement was a quiz question. *Grouping* can be used to specify activities with an indirect relation to the object activity of a statement. For instance, a qualification has an indirect relation to a class and can therefore be specified using the *grouping* structure. *Category* is used to add activities that can categorize/tag a statement. The only example given in the xAPI specification is that the xAPI profile used when generating statements can be specified using *category*. The context type *other* can be used to specify activities that are not found to be appropriate in any of the *parent*, *grouping*, or *category* context types. The example given in the xAPI specification is that an actor studies a textbook for an exam, where the exam is stated to belong to the context type *other*.

Extensions, like ContextActivities, are organized in maps. They should include domain-specific information that is not covered using the other context properties. The map keys for extensions must be represented in Internationalized Resource Identifier (IRI)s; the map values can be any valid JSON data structure such as string, array, and object. Thus, using extensions to express context information in an xAPI statement is more flexible than using ContextActivities. As such, the advice in the specification, regarding interoperability, is that built-in xAPI elements should be preferred to extensions for storing information, if available (Advanced Distributed Learning, 2017c). The xAPI specification gives the example of an actor using a flight simulator, where altitude, wind, and GPS coordinates can be expressed using extensions.

Since xAPI allows for registration of such a diversity of context-related information, through both the context structure and the result structure, data described in xAPI may for instance be used for personalization, visualization, assessment, and prediction.

Comparing the different context data models to xAPI

Having examined context in xAPI, we now look at its similarities and differences with regard to the previous research on context data models (see Table 2).

xAPI collects data regarding the learners/agents and their activities. Of the context data models presented above, all except the LOCO model also have the learner (or user) as their unit of focus. LOCO, however, focuses on learning objects (in xAPI, learning object information would be represented in the object of the xAPI statement, rather than the context). In terms of flexibility, xAPI is quite flexible regarding data registration, similar to CAM. The other data models appear generally to be more rigid, for example due to stricter specification of available properties and data types.

Table 2 Context data model and xAPI comparison

Name	Unit of focus	Flexibility of data registration	Categorization of context	Interoperability	Usage
LOCO	Learning object	More rigid	No categorization	Targets interoperability	Personalization, examine object use
CAM	Learner	Flexible	No categorization	Targets interoperability	Personalization, visualization, examine object use
LCDM	Learner	More rigid	Two-level categorization (high/low level)	Targets interoperability	Visualization
RCM	Learner	More rigid	Two-level categorization (high/low level)	Does not target interoperability	Personalization/contextualization, visualization
xAPI	Learner/agent	Flexible	No categorization	Targets interoperability	Personalization, visualization, examine object use, assessment, prediction, etc.

Concerning classification of context, the LCDM has capabilities for two-level categorization of events, making a distinction between high-level and low-level categories, for example the high-level categorization "environment" and the low-level categorization "noise level." The RCM approach, similar to LCDM, suggests both high-level categorizations of context (in the form of context dimensions) and low-level categorization (information belonging to the separate context dimensions). In contrast, xAPI and the other data models do not provide this type of classification of context.

Interoperability can be enabled in varying degrees through usage of common data models/specifications, depending on how they are used/defined. As such, interoperability is a stated end for all the data models, except RCM. While the RCM approach is used for data analysis with regard to personalization and contextualization, this approach does not specifically target interoperability. Instead of using a standardized approach, such as requiring terms to be chosen from pre-established vocabularies when generalizing the RCM to a new application domain, the configuration is done at an ad hoc basis for each domain added (e.g., for each new domain, the data format must be specified). Concerning the use of the context data models, there seems to be an emphasis on personalization (e.g., providing recommendations) and visualization. CAM and LOCO also examine the use of learning/data objects. While data from xAPI may be used for personalization, visualization, and examining object use[4], it provides structures for describing data that cannot be described with the context data models (e.g., information about learner results). Thus, provided the xAPI data are sufficiently described, they may also be used for other purposes, such as assessment and prediction.

AVT project—the case

The case study examined the use of xAPI in the AVT project (Morlandstø et al., 2019; Wasson, Morlandstø, & Hansen, 2019), which ran from August 2017 to May 2019. We chose AVT as a case study subject due to it being a real-world project that used a learning activity data standard (i.e., xAPI) for describing data originating from multiple sources, thereby having the potential to uncover challenges and

[4]Data related to learning object use are typically stored in the object (activity) of an xAPI statement.

limitations of data description and integration as they unfold in practice. AVT, owned and funded by the Norwegian Association of Local and Regional Authorities (KS), was initiated by the Educational Authority in the Municipality of Oslo (Utdanningsetaten), and the Centre for the Science of Learning & Technology (SLATE), University of Bergen was responsible for research and for leading the project. In addition, the project group comprised 9 vendors from the Norwegian EdTech sector, and 4 of the schools in Oslo. The project consulted representatives of the Learning Committee (Læringskomiteen SN/K 186) under Standards Norway, the organization that is responsible for the majority of standardization work in Norway (Standards Norway, 2019) and representatives from The Norwegian Directorate for Education and Training (Utdanningsdirektoratet), as well as taking feedback from the Norwegian Data Protection Authority (Datatilsynet), the Norwegian Competition Authority (Konkurransetilsynet), and representatives from the Parent organization for schools (Foreldreutvalget for grunnskolen) and the Student organization (Elevorganisasjonen).

The AVT project explored possibilities for using activity data to adapt learning to individual learner needs, and for formative and summative assessment at the K–12 level. Since learners generate activity data in a number of tools from different vendors, a challenge is how to integrate such data, to provide richer information on the activities of each individual learner. Therefore, AVT looked at data sharing among different EdTech vendors, resulting in the implementation of a framework that helped to standardize data originating from different educational tools and systems, and which enables secure data flow among vendors. xAPI was the chosen format for data description, integration, and data exchange. To enable more consistent use of xAPI, the project used a number of concepts from a vocabulary that had been adapted and translated to Norwegian by Læringskomiteen SN/K 186, as they are working with learning technology and e-learning (Standards Norway, 2020). SN/K 186 also participates in projects that develop artifacts based on standardization initiatives, such as AVT.

Method

This research adopted an exploratory case study methodology (Oates, 2006), see Fig. 2, using AVT as a real-world case. The subject to investigate was challenges and limitations of using a current learning activity data standard (i.e., xAPI) for describing learning context with regard to interoperability and data integration, and how these might be addressed. Consequently, we involved stakeholders at the different stages of the research process. This paper addresses the first four steps of Fig. 2.

Initially, we studied the AVT project documents, prepared interview guides, did sampling and recruitment of participants, and prepared consent forms. Next, we interviewed stakeholders from the AVT project about the gaps and needs of xAPI, with emphasis on descriptions of context. To provide important background information, we also asked about the rationale for choosing xAPI and the process of using xAPI for data integration and data sharing. Using thematic data analysis, we then identified themes emerging from the interview data. Subsequently, based on interview data and inspection of the xAPI and xAPI profile specifications, we formulated recommendations for how xAPI context can be improved regarding interoperability and data

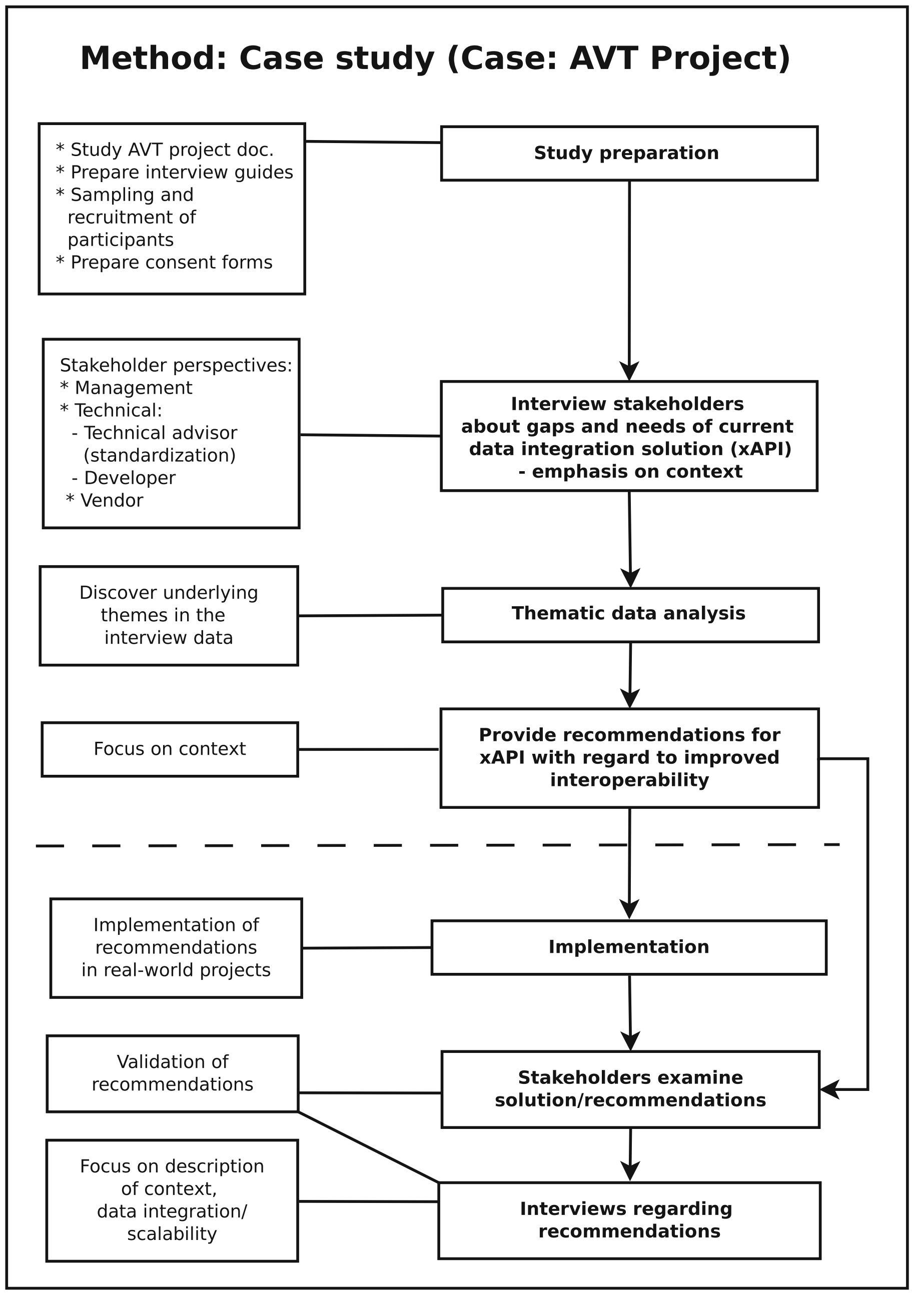

Fig. 2 Research methodology

integration. In this paper, we provide a summary of the recommendations. In future work, we will provide a detailed account of the recommendations, implement a number of the recommendations in two separate projects, and validate the recommendations through stakeholder examination and interviews.

Participants

The selection of participants was done through purposive sampling (Bryman, 2012, p. 418). Purposive sampling of participants is not done at random, but rather in a strategic way (Bryman, 2012). The point is to select participants that are appropriate for the research questions. It may be important with variety in the sample, meaning that the sample members differ in their characteristics as they are relevant to the research questions.

Eight stakeholders in the AVT project were recruited for the interviews to identify gaps and needs of xAPI (see Table 3). The first seven interviews were conducted between October 22 and November 01, 2019. Through these seven interviews, it became clear that our sample was missing an important AVT member related to the questions we wished to answer in our research, thus an additional interview was conducted on February 07, 2020.

The participants worked on a diverse set of tasks within the AVT project and their roles represented different perspectives. Two participants represented a developer perspective (i.e., they had experience with implementation of xAPI methods and preparation of datasets), three participants represented a leader perspective (two related to decision making for AVT; one was the leader of an external organization associated with AVT), two participants had a vendor perspective (one of these vendors had delivered data to AVT and the other had not), and there were also two technical advisors in the sample. The advisors were knowledgeable regarding standardization within the educational domain and gave advice to the rest of the project about how to use xAPI for describing activity data and context; they also made some examples of xAPI statements that describe activity data related to AVT, which the developers subsequently followed/used as a template.

Data collection

Data collection was conducted using semi-structured interviews where the objective was to find answers to the following overarching questions:

Table 3 Participants, sorted by interview order

Identifier	Gender	Perspective	Tasks (sample)
P1	Female	Leader	Conducting meetings, delivery of documents, some technical work
P2	Male	Developer	Technical tasks within AVT (e.g., server, database, data sharing, security), contributed to xAPI example statements
P3	Male	Leader	School owner representative, specifying how vendors should code xAPI activity data
P4	Male	Technical advisor	Vocabulary/profile work, detailed work on how to represent context for AVT activity data, created xAPI example statements
P5	Male	Vendor (delivered data), developer	Planning and implementation of vendor solution
P6	Female	Vendor (did not deliver data)	Project leadership and coordination for vendor
P7	Female	Leader (external organization associated with AVT)	Conducted meetings where a number of AVT members participated, which fed into the AVT project; work related to vocabularies and their use in Norway
P8	Male	Technical advisor	Vocabulary/profile work, detailed work on how to represent context for AVT activity data, participated in developing xAPI example statements, explored and informed vendors about tools and libraries for storage and exchange of activity data

– What was the rationale for choosing xAPI in the AVT project?
– What was the process of using xAPI for data integration and data sharing?
– What challenges and limitations of xAPI were identified when describing context for data integration?

Interview guides were developed based on study of several documents including the final report for the AVT project (Morlandstø et al., 2019), the xAPI specification (Advanced Distributed Learning, 2017a), and the xAPI profile specification (Advanced Distributed Learning, 2018a). The interview guides contained a list of topics to be covered in the individual interviews, which can be broadly be categorized as the following:

- Background (e.g., regarding the role/tasks of the participant in the AVT project, and their previous experience with xAPI),
- Leadership (e.g., related to reasons for choosing xAPI for the AVT project, and other decisions made within the project),
- Technical development (e.g., practical experiences of describing data in xAPI),
- Context (details about how context was represented using xAPI and reasons),
- High-level topics (e.g., benefits and challenges of using learning activity data specifications for data integration).

The participants were interviewed according to their perspectives, roles, the tasks they had worked on, and their areas of competence. Each topic in the interview guide was covered by at least two participants. Before the interviews started all participants were presented with a consent form, explaining aspects such as the purpose of the project, that audio would be recorded for subsequent transcription, that participant information would be anonymized upon transcription and stored securely, and that their participation was voluntary and could be withdrawn at any time. Because the audio recordings could theoretically be used to identify the participants, the project was reported to the Norwegian Centre for Research Data (2020), which approved the project based on the measures taken concerning privacy and research ethics. Participants were informed that the interviews could take up to 90 min, although most interviews finished in less than an hour. All interviews were conducted in Norwegian.

Data analysis

The study used a thematic analysis approach (Bryman, 2012) to analyze the transcribed interview data. The transcripts were collated and read through several times for familiarity with the content. The interview data were coded at two levels, using NVivo (2020). First, the data were coded according to our overarching questions; next, the data for each overarching question were coded at a more fine-grained level and themes were identified through an inductive process.

Findings

The analysis resulted in 32 codes during the second level of coding that were further aggregated to seven themes, each pertaining to an overarching question (see Fig. 3).

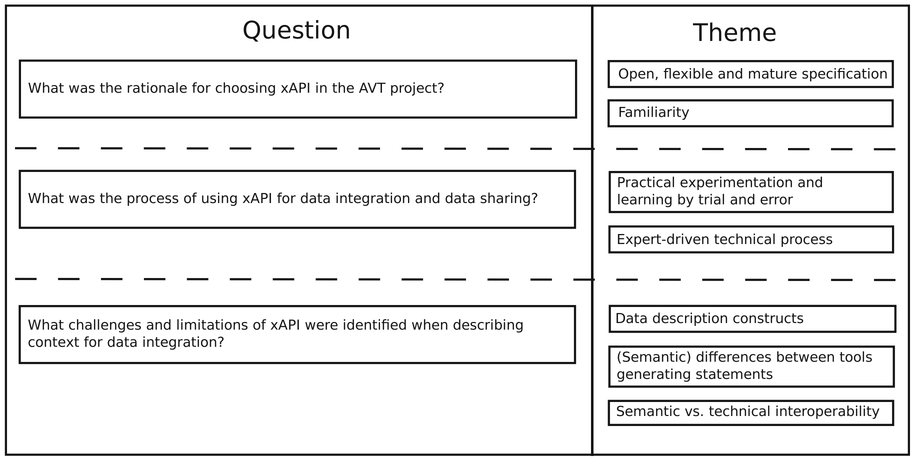

Fig. 3 Questions and themes

The findings in each theme are summarized below. All quotations in this section have been translated from Norwegian to English.

Rationale for choosing xAPI

Open, flexible, and mature specification

A number of the responses given by the participants indicated that xAPI was chosen by AVT because it was open, flexible, and mature. P1 and P3 (representing a leadership perspective) explained that at the time of making the decision, xAPI had already been chosen as the preferred learning activity data specification for Norway by Standards Norway and the SN/K 186 committee, and this was the main reason that AVT chose to use xAPI. Another reason, mentioned by P1, was the openness and flexibility of xAPI. Since the choice of xAPI by SN/K 186 was the main reason that AVT decided on xAPI, the rationale for choosing xAPI by SN/K 186 was also of interest during the interviews.

As P3, P7, and P8 had actively taken part in or observed the choice of xAPI by Standards Norway and SN/K 186, they explained reasons for the choice by the committee. Openness and flexibility were mentioned by all three participants. P3 and P7 specifically remarked on the possibility to customize xAPI for a particular use-case/project (e.g., through profiles or extensions). P8 stated that while IMS Caliper was an alternative at the time of the SN/K 186 decision, it seemed less mature than xAPI. Particularly, xAPI had more extensive documentation than Caliper. Two other aspects for choosing xAPI over Caliper, mentioned by P8, was that Caliper is more adjusted to the US educational system and that vendors of the big EdTech systems that have the greatest influence on Caliper. Looking at the IMS members, it is clear that the majority are US companies and institutions (IMS Global, 2020). Griffiths and Hoel (2016) confirm that it is the member organizations of IMS that influence the use cases that Caliper can describe and that the specification seems to target the needs of larger vendors and institutions. Interestingly, P7 explained that SN/K 186 had not taken a definite stand that they would only use xAPI, but they would try out the specification.

Familiarity

Another reason for choosing xAPI for the AVT project, as revealed by P1, was that several project members were already familiar with the specification. P4 and P5 confirmed that they had both used the specification in their work for EdTech vendors (P4 functioned as an advisor in the AVT project). The AVT project members also had some knowledge about other Norwegian users of xAPI. P7 and P8 were aware of a smaller EdTech vendor that was using xAPI. P7 mentioned that a large higher education institution in Norway had experimented with the specification.

Process of using xAPI for data integration and data sharing

Practical experimentation and learning by trial and error

Using xAPI for data integration was very much a process of practical experimentation and trial and error, as P3 explained. As a starting point, the project used data in xAPI format from math tests in Oslo municipality (P2 converted the data to xAPI format). Having access to these data, a group including the technical advisors (P4 and P8) and P2 made one simple and one advanced example statement (more examples were later added), as explained by P3.

P1 and P3 stated that the vendors were asked to use the examples as templates or rules for how to construct xAPI statements. The examples used concepts (e.g., verbs) from the xAPI vocabulary, i.e., the collective vocabulary defined in the published xAPI profiles (Advanced Distributed Learning, 2020b), some of which had been translated to Norwegian by SN/K 186. Other concepts used in the examples, which were not available in the xAPI vocabulary, were defined in a separate xAPI profile. According to P4, following the examples was to ensure more uniform data descriptions, thus making data more easily integrable. For storage of xAPI statements, vendors were encouraged to implement their own LRS, which would accept queries in a specific format and return statements.

Concerning data sharing, a prototype was developed that could be queried for student data. The prototype had an LRS component (storing data from the Oslo tests), access control (to allow secure data sharing from other LRSs, e.g., from vendors), and a limited user interface that could display some data about students, as explained by P2.

Initially, a number of the participating vendors in the AVT project indicated a willingness to share data. To make it easier for vendors to share data, thus to get more data for the project, the requirements for how statements should be formulated were eventually eased, as explained by P1. While P6's company did not manage to deliver data in time, she did state that this decision would have allowed her company to deliver data eventually, as the data generated by their tool was at a higher level compared to the xAPI examples they were given. Also, P5's company, who did deliver data, ended up delivering data in a format that differed slightly from the examples. This was due to internals of their application since some data generated did not fit into the built-in structures of xAPI. As P1 explained about the benefits and drawbacks about easing the requirements for the statements: "It gives advantages since we might get some additional data, but it gives drawbacks related to subsequent analysis. Because the data consistency, i.e., the quality of the entire data set, may not be as good. So, we have to weigh [drawbacks and benefits of] this the entire time." Still, at the end of the AVT project, only one vendor managed to share data in xAPI format. The data shared by the vendor was technically integrated with the Oslo municipality test data, but there was no individual student whose data appeared in both data sources (as stated by P1). Thus, more could have been said about data integration if more data would had been successfully delivered.

Expert-driven technical process

The AVT work on using xAPI for data integration and data sharing was very much driven by the technical advisors (P4 and P8) who were knowledgeable about standardization and xAPI. When asked in-depth about how the project chose to describe context using xAPI, and why that decision was made, those participants who had a leadership perspective in AVT (P1 and P3), and P2 (developer), all said to rather ask the advisors P4 and P8. As P3 put it: "P4 and P8 are the two masterminds behind the very technical parts of this [project]." He stated that the project relied heavily on P4 and P8 in the area of modelling context and that the other AVT members trusted their recommendations. P4 and P8 contributed greatly to the xAPI profile and to the example xAPI statements. In addition, P8 worked on specifying vendor LRS requirements. When asked if it was clear how to construct xAPI statements in terms of data integration, P4 stated that he had hoped more vendors would have shown an interest in how the statement examples were formulated, as this could have sparked interesting discussions. From the perspective of a vendor who delivered data, this view of vendors as showing a limited interest in statement formulation was confirmed by P5. He expressed that how statements were generated to allow for data integration from multiple sources was not a focus area for his company: "As a vendor this is not something we care about. Rather, it is those who consume our data that have to figure this out."

Challenges and limitations of xAPI

Data description constructs

In terms of describing context within the AVT project, concepts from the xAPI vocabulary were used. While this approach enabled the description of the data, P3 noted that the example xAPI statements did not use many of the concepts from the xAPI vocabulary, as correct use of the concepts was quite challenging. For instance, for the verb "answered," the many different types of items that could be solved or answered by a student could make it challenging to properly define the item. To add further concepts related to context that were not available in the xAPI vocabulary, an xAPI profile was developed for the AVT project, as explained by P3. The concepts were registered as metadata in the form of activity types. The activity types were used to represent concepts relevant for the AVT project, such as competence objective, school owner, and school. Thus, instances (i.e., activities) of these activity types could also be part of the statements. As P4 and P8 pointed out, using a profile allowed for restriction of the concepts that were used in xAPI statements. P8 also emphasized that profiles can explain the specific meaning of a concept.

Related to data description, two participants (P3 and P5) mentioned that mapping real-world data to the xAPI format could be difficult, due to the assumptions of xAPI. P5 mentioned a problem of registering the answer to an assignment in xAPI, where xAPI expected exactly one answer. The solution of the vendor that P5 represented, however, had multiple choice. Therefore, more than one answer could be correct (in combination).

To add contextual data to the context structure of the xAPI statements, in addition to the data that could be added to the seven properties taking a single value or object that represents a single entity, P4 and P8 explained that the project had the choice between adding them to ContextActivities and extensions. In the end, it was decided to use ContextActivities, with the following structure: context -> ContextActivities -> *grouping*. Accordingly, activities which collectively constituted the context data were grouped together. P4 said that part of the reason was that the example statements were inspired by CMI-5, a specification that has as its goal to enable a consistent information model for xAPI (CMI-5, 2020). He also stated that ContextActivities is a more standard way of using the xAPI specification, while extensions are a more custom approach. As P4 explained regarding extensions: “No one can know anything about how that format is.” While P4 was initially in doubt about the use of ContextActivities, he did feel it was a better choice than using extensions.

P8 pointed out that while ContextActivities had a fixed structure, and extensions had a more flexible structure (e.g., it can contain a JSON object or a list). He stated two reasons for choosing ContextActivities over extensions: The first was related to the libraries available for generating xAPI statements. It was much simpler to add ContextActivities than extensions. The second reason was that vendors were to set up their own LRS according to the xAPI specification. One requirement was to set up a query API, so that the LRS could be queried in a standardized way. As P2 stated, however, it was a challenge that LRSs do not have good query capabilities, the focus seems to rather be on data. Using ContextActivities made it easier to filter statements on contextual data (activity types). P8 explained that to filter based on the contextual data when using extensions, a new API would need to be built on top of the query API: “You would have to build another API on top of the existing, to enable those queries. So, in a way it [using ContextActivities] is a minor ‘hack’(…). It was a way to enable more powerful queries or search.” Thus, using ContextActivities enabled enhanced query capabilities without requiring extra development work. P8 pointed out that using ContextActivities to group contextual information had a significant drawback in terms of semantics. The activity types added to the xAPI profile, such as school, school owner, and competence objective, were not really activity types. Thus, the activities added to ContextActivities, which were instances of the activity types added in the AVT xAPI profile, were not really activities. In a report on xAPI, Learning Pool, the developers of the Learning Locker LRS, tell a related story on how not adhering to semantics can make querying easier. They wanted to model that a user had liked a specific comment by another user (Betts & Smith, 2019). Semantically, the comment (an activity) was the object of such a statement, and the other user belonged in the context. In terms of querying, they wanted to count how many times a user had made comments that were liked by other users. In this instance, they found querying easier if the other user was in the object structure, and the comment was in the xAPI context, even if this was not semantically correct. They commented: “This is just one example of how adopting the specification has been somewhat harder than we first thought it would be.”

Another aspect of ContextActivies is how contextual information is grouped. When asked why *grouping* was used in place of *category* or *other*, both P4 and P8 were uncertain. P4 stated he had read a lot about the topic, but he was unsure about how they differed. There seemed to be a lack of semantics that set them apart. P4 said any one of them could be used in the statements, or they could all be used: "We could have said that competence objective is in *grouping*, but the school owner, which school it originated from, is in [*category*] or [*other*]. But that would just make the statements even more complex." In the end, P4 said they just needed to make a rule for which one to use, to ensure all vendors would follow the same procedure (and to not confuse the vendors). P8 also pointed out the difficulty of choosing between *grouping*, *category*, and *other*. When reflecting on the differences he stated: "*Grouping*… I don't really know what that means, really." He added that choosing between the three was made even more difficult since the activity types of the added activities were not really activity types.

(Semantic) differences between tools generating statements

Another challenge, expressed by P3, is that the openness of xAPI allows for the combination of data that may look similar, but that turn out to be not comparable. P4 mentioned duration as an example of a property that may be used differently in different EdTech systems and thus could be difficult to compare across sources. A property that was in fact problematic in the AVT project is the score of a student when solving a specific item, as mentioned by P3 and P5. Even when a score is set according to the same scale, it may not have the same meaning across systems. In AVT, the meaning of score was different in the tests administered by Oslo municipality and in the data delivered by the vendor, as their tool was a system for practicing, rather than testing. P4 and P5 note that identifying these types of discrepancies require the sources to be well-known by those using them (e.g., for analysis). Because of the xAPI openness, different tools may also generate a different number of statements for the same type of event, as pointed out by P3. He stated that this was not a big challenge for the AVT project, because only one vendor delivered data. When more than one vendor delivers data, however, it could become a challenge. Another challenge, concerning tool differences, is that the same type of data may be recorded at different levels of granularity by different tools. For instance, P6 expressed that her company attempted to deliver data that was at a higher level compared to the xAPI examples they were first given.

Semantic vs. technical interoperability

The process of using xAPI for data integration was expert-driven, as mentioned earlier. This was particularly apparent when asking the different stakeholders about data descriptions. When asking those with a leader perspective in AVT (P1 and P3) and those with a developer perspective (P2 and P5) if xAPI could satisfactorily describe the data, they generally agreed. As P3 stated: "There is no data we have wanted to describe that we have not been able to describe with xAPI so far." When probing further about representation of context in xAPI, P1, P3, and P5 generally agreed that xAPI could do it in a satisfactory way. The developers based their data descriptions on the examples created by P4 and P8; thus, they knew xAPI more on a technical level of interoperability than on a semantic level. As P2 expressed it: "I didn't examine other options for xAPI. I just thought that here we have an example, then I will fill out the data I have according to the example." When P3, representing the leadership perspective in AVT, was asked if he thought it would have been more difficult to describe the data without the example, he agreed: "Yes, you can express almost anything with xAPI, so you have to start with a need." When asking the technical advisors, who were more knowledgeable about xAPI and standardization, about data descriptions and context, they both mentioned several problems related to the openness/flexibility of xAPI and semantics. The different ways to represent context data (ContextActivities -> *grouping*, *category* or *other*) was one example. They also saw problems in constructing xAPI statements in terms of data integration because of the many ways data of the same type might be described. P4 said that in the context of AVT you need to try to find concepts appropriate for the vendors and to make them follow rules/templates. P8 said creating statements for data integration might be achievable within a small project such as AVT, based on common rules and documentation. If combining data with another community of practice, however, it would be challenging: "There would be many sources of error concerning syntax, semantics, etc." He suggested library development and schema validation as possible ways to alleviate this challenge.

Recommendations

We see that the analysis of the stakeholder interviews establishes that there is a lack of clarity in how to describe xAPI context data related to interoperability and data integration, which negatively affects the expressibility of xAPI. The challenges identified through systematic analysis of interview data and inspection of the xAPI and xAPI profile specifications are shown in Table 4.

Based on the challenges identified, we provide recommendations (i.e., recommended solutions), in summarized form, on how xAPI can be improved to support interoperability and data integration, with emphasis on descriptions of xAPI context (see Table 4). The recommendations were specified using an iterative process, where we used a bottom-up approach of identifying challenges in the analyzed data, xAPI and xAPI profile specifications, and a top-down approach of examining relevant research literature on context categorizations. Although the emphasis has been on descriptions of context, we additionally give some recommendations that pertain to the xAPI framework as a whole (e.g., data typing and validation, and improved documentation), as they are important with regard to the expressibility of xAPI context. In our recommended improvements to xAPI, it should be noted that some recommended changes will require a change to the xAPI standard (or documentation), while others can be implemented using an xAPI profile. For this paper, the recommendations are at a conceptual rather than implementation level.

Regarding the recommendations, data typing and validation for specific use cases can currently be addressed through two xAPI profile constructs (Advanced Distributed Learning, 2018b). Statement templates describe a way to structure xAPI statements, and can include rules, e.g., for restrictions on the data type of a specific property value, while patterns define how a group of statements should be ordered. Both constructs can be checked by an xAPI profile validator.

Recently (August 2020), Advanced Distributed Learning (2020a) published information that there are plans to standardize xAPI 2.0, which is an upgrade from the current version 1.0.3 (Advanced Distributed Learning, 2020a). An IEEE LTSC working group,

Table 4 Challenges identified and recommended solutions

Category	Challenges identified	Recommended solutions
Context	The distinction between ContextActivities and extensions appears artificial, and it is not always clear which to use. Extensions are flexible in how data can be registered and therefore could make it more difficult to integrate data. ContextActivities is not a good fit for all types of context data.	Use a unified structure for context in xAPI, i.e., context dimensions, with appropriate (low-level) properties for each context dimension. Use data typing and validation to restrict properties and value types for context dimensions. Depending on the property, the value type can be Activity, but other value types should also be supported (e.g., JSON object and string).
	Grouping of related activities in ContextActivities can be done within three different structures (*grouping*, *category* or *other*). It is not clear how the three structures differ. The grouping structures are all very high-level.	Remove distinction between *grouping*, *category*, and *other*. All context data that do not have an explicit (parent) relationship to the statement should be placed in the same structure. For the suggested unified structure for context, i.e., context dimensions, the related information can be listed more explicitly as property values belonging to the appropriate dimension and property.
	The query capabilities of LRSs are seen as limited. The example given is that it is only possible to filter statements on contextual data that are instances of an activity type. Thus, to allow filtering of resources in AVT (without extra development work), resources were registered as activities, even if they were not really activities on the semantic level.	The xAPI specification defines the query interface that all LRSs must implement (Advanced Distributed Learning, 2017d). In the case of filtering based on resources, the specification needs to be extended, so that it is possible to filter contextual data by any resource type. Individual LRS providers have addressed this issue on an ad hoc basis (Learning Locker, 2020), but the problem needs to be further addressed in the xAPI specification to ensure LRS interoperability (e.g., for xAPI users not to have to rewrite substantial amounts of code if moving data between LRSs).
	Concepts from the xAPI vocabulary not sufficient to describe all data.	Use of xAPI profiles to add additional concepts.
	The same vocabulary concept may be represented in different public xAPI profiles, which make up the xAPI vocabulary.	Stricter curation/approval process of the public profiles.
Data typing and validation	Tools generating data at multiple levels of granularity is a challenge, which may make it more difficult to meaningfully integrate data.	To help tool developers identify and enforce the expected level of granularity, xAPI data typing and validation can be used. For instance, if a property takes a list of activities (more granular), validation can ensure that less granular values (e.g., integer) will not be accepted.
Data typing and validation	Difficulties in mapping real-world data to xAPI due to its assumptions.	Data typing and validation can help to ensure that the assumptions of xAPI (e.g., expected value for an xAPI concept) are made more explicit and tested against the data, to avoid wrong use of the specification. It is also crucial that the xAPI specification can be extended as new use cases reveal new needs for data registration.
Data typing and validation	Different tools may generate different numbers of statements for the same type of event.	Validation could be tied to the number of statements generated for a given type of event, and to ensure that the statements generated follow an ordered pattern.
Data typing and validation Documentation	The openness and flexibility of xAPI allows data and relationships of the same type to be modelled in a myriad of different ways.	Add clearer modelling guidelines to the documentation. Add data typing and validation of properties and property values. Use of profiles to specify vocabularies.
Documentation	It may be challenging to correctly use concepts from the xAPI vocabulary.	Improve xAPI documentation, e.g., document more solutions for specific use-cases, and add more examples of how to use the xAPI vocabulary concepts in order to avoid misunderstandings and remove ambiguity.

comprising stakeholders from the xAPI community and technical experts, have agreed on the new standard; however, a formal balloting process also must be conducted in order to standardize. While there is information that some new structures will be introduced in order to describe context, it is indicated that these structures will allow more structured descriptions of individuals (i.e., contextAgents) and teams (i.e., contextGroups); thus, we do not believe they will solve the issues/challenges related to context that our research has identified and for which we provide recommendations in this paper. Another addition to the proposed xAPI 2.0 standard is a best-practices guide, "which will be linked to the eventual standard as a living document that can grow and change with advances in learning science and technologies" (Advanced Distributed Learning, 2020a). Based on the published information, this guide could help in the identified need for improvements in the xAPI documentation. Other changes in the proposed version 2.0 include forbidding additional properties in statements and standardizing timestamps. While the changes may help in terms of interoperability and data integration, they do not specifically relate to the challenges we have identified and the recommendations we have provided.

Discussion

Two research questions were posed in this paper, regarding (1) gaps and needs of xAPI in terms of interoperability and data integration focusing on context descriptions, and (2) how identified gaps and needs can be addressed in order to provide improved interoperability and data integration. We have addressed RQ1 through analysis of the data from the AVT stakeholder interviews and inspection of the xAPI and xAPI profiles specifications, and RQ2 through providing summarized recommendations on how xAPI can be improved to support interoperability and data integration with emphasis on descriptions of xAPI context. In the following, we discuss patterns and trends related to xAPI data descriptions and interoperability/data integration, which we have identified based on the review of research papers that utilize xAPI to describe data.

Although papers on xAPI commonly mention the benefits of xAPI in terms of interoperability, most of the studies that use or explore the use of xAPI to describe data worked with only one data source (Hruska, Long, Amburn, Kilcullen, & Poeppelman, 2014; Megliola, De Vito, Sanguini, Wild, & Lefrere, 2014; Papadokostaki, Panagiotakis, Vassilakis, & Malamos, 2017; Wu, Guo, & Zhu, 2020; Zapata-Rivera & Petrie, 2018). In such cases, there is no practical experience with the challenges and limitations regarding data descriptions and interoperability. For instance, the challenges related to the flexibility of xAPI and data descriptions do not readily appear. The data can be described in different ways, all accepted according to xAPI. It is when trying to integrate xAPI data from different sources the challenges of inconsistent descriptions will surface. Thus, out of the five referenced examples using only one data source, four of them do not touch on challenges related to xAPI and data descriptions/interoperability. The exception is Hruska et al. (2014) who examine challenges with, and give examples on, how to encode information about teams/groups in xAPI statements, including descriptions of group context. Interestingly, Megliola et al. (2014) have many reflections on vocabulary (verbs and objects) for describing events in their modelling domain (aeronautics), but the reflections are based on multiple theories in linguistics rather than the practical application in xAPI.

Having examined the knowledge base, we have found a very limited number of studies that utilize xAPI to integrate data originating from multiple data sources and share lessons learnt from such a project. The CLA toolkit case study (Bakharia et al., 2016) was one such study. Here, data from different social media were described in xAPI format and integrated in an LRS for use in a systemic analytics solution. The project included designing a common vocabulary through re-use of concepts from W3C ActivityStreams 1.0[5] and providing mappings from individual tool concepts to the common concepts. They also examined how context could potentially be described and made decisions on how to describe context in the project. At the time of the study, xAPI profiles had not been added to xAPI. Thus, vocabulary and data type/validation rules could not be described in a machine-readable manner. Rather, the vocabulary, together with prescriptions on how to describe social media data, were stored in a recipe. A recipe is a textual description on how xAPI statements for a certain experience can be constructed (Miller, 2018). At the time, recipes were the common means to share the implementation of an information model with a community of practice, serving as a potential aid in terms of interoperability. They were, however, not machine-readable. Through actual usage of xAPI, the researchers working on the CLA toolkit were able to identify challenges and complexities of the data integration approach. Among the lessons learnt were that providing xAPI context data, while optional according to the standard, was essential for their project. In addition, they recommended that xAPI be extended with the JSON-LD specification, as the lack of machine-readable vocabularies and rules were a weakness of the xAPI specification.

Following the research by Bakharia et al. (2016), we see that xAPI has introduced capabilities of machine-readable vocabularies and rules, since the xAPI profiles specification has been added to xAPI. Thus, machine-readability is no longer a core problem of xAPI (although it is of importance that tools and libraries implement the methods needed to read metadata descriptions and apply data typing/validation rules). It is encouraging to see that xAPI has used the results from research when choosing to add JSON-LD capabilities.

xAPI, through xAPI profiles which are defined using the JSON-LD format, leverage semantic technologies to allow for documents that are not only readable by humans, but also machines. In their article, Verborgh and Vander Sande (2020) discuss the importance of not conducting research related to semantic technologies in a vacuum (e.g., controlled research experiment). While researchers are often reluctant to take their solutions out of the labs, deeming large-scale technology use as a trivial engineering problem, this article highlights that practical use of the technologies outside of the safe research environments, e.g., through integrating data from sources containing real-world rather than synthetic data, is likely to uncover new challenges that need to be solved in order to promote adoption among practitioners. The results from our interviews indicate that there is indeed a need for more practical research with real-world data description and integration.

Research has shown that data integration within LA, an important part of LA scalability, is a challenge in itself. Previous research in the domain of LA and higher education has found that if data are integrated, they typically originate only from a few data

[5]https://activitystrea.ms/

sources; when data are integrated, the integration is often of data of similar formats (technically, this type of integration is easier than combining data of different formats), and there seems to be little use of learning activity data specifications such as xAPI (Samuelsen et al., 2019). One reason that xAPI is not used more for data integration may be that many tools do not support providing data in the form of xAPI statements. In such cases, the use of xAPI can be challenging, since transforming data to xAPI format is likely to require a considerable effort. Depending on the situation, it may be more convenient to store data of similar formats originating from different data sources in a NoSQL database (e.g., if the data integrated are provided as JSON through REST APIs), even though the problem of alignment of concepts from different sources will still be a challenge. In some cases, the integration approach may also be a manual one, e.g., through copying and pasting data from data sources into excel sheets; an approach that may also require considerable human effort. In the case of xAPI and other learning activity data specifications/standards, there seems to be a need for more tools supporting their formats.

While previous research has found little evidence of xAPI use in the research literature (Samuelsen et al., 2019), the xAPI guide by Learning Pool provides information that there are tools that can export data in xAPI format (Betts & Smith, 2019). These include LMSs, content management systems, and authoring tools. The tools often provide the xAPI export functionality through plugins. Due to the flexibility of xAPI data descriptions, however, one cannot expect statements from these tools to meaningfully integrate (i.e., to scale), since different plugins, made by different developers, may use different syntactic and semantic structures for describing the same data. For the statements of an xAPI-enabled tool to be compatible with statements from a number of other xAPI-enabled tools, we can imagine that the tool will need a number of similar plugins, each plugin making statements compatible with statements from a specific tool or set of tools. While there are architectures (Apereo, 2020; JISC, 2020) that provide connectors/plugins for several tools to store their data as xAPI in a common data store, the problem is still not solved since different architectures can also model data of similar types in different ways. Thus, there is a need for tools that can provide xAPI data in a coherent format.

Looking at studies using xAPI, we find, consistent with its flexibility, a diverse set of examples of how researchers add, or intend to add, context data. Not all examples seem to be within the intended use of xAPI. Wu et al. (2020) demonstrate using ContextActivities—*parent* to store a verb as an activity, even though a verb is not an activity, and should not be the parent of an activity. Sottilare, Long, and Goldberg (2017) propose to add course to the xAPI result structure, even though it is added in the published xAPI vocabulary as an activity type that can be used in the context structure with ContextActivities (Advanced Distributed Learning, 2020b). We also find examples related to the different ways ContextActivities can be used to describe data. Several works use ContextActivities—*grouping* and ContextActivities—*other* in the same statement (Claggett, 2018; Hruska et al., 2014), even though we have identified that it is not really clear that there is a difference between the two structures. These examples confirm our findings that there is a need to enhance the expressibility of context in xAPI. The recommended addition of context dimensions would be one way to enable more consistent description of the data. Documentation, data typing, and validation should provide clarity and guidance regarding how to achieve this goal.

In the xAPI specification, learning objects and their context are represented through activities. An activity has an id (an identifying IRI), and an optional definition stored in an Activity Definition Object (Advanced Distributed Learning, 2017c). The activity definition can be stored together with the activity id in the xAPI statement, or it can be stored in the activity id IRI and downloaded by the LRS (hosted metadata solution). While xAPI can represent learning objects and their context, the topic of representing learning objects has also been addressed by several other standards, including "ISO/IEC 19788 Information technology - Learning, education and training - Metadata for learning resources" (International Organization for Standardization, 2011) and "Encoded Archival Context – Corporate bodies, Persons and Families" (Mazzini & Ricci, 2011). Thus, the xAPI community could look to such standards when aiming to provide future improvements for learning object representation.

The challenges we have identified with xAPI may lead us to question if IMS Caliper would currently be better suited for representing interaction data in the educational domain. Examining the specification (IMS Global, 2018), we find there are 14 available profiles (called metric profiles), which target experiences such as assessment, forum, and tool use. Similar to xAPI, Caliper also supports machine-readable data descriptions through JSON-LD. It seems that Caliper has the potential of avoiding many of the xAPI problems caused by flexibility. For instance, it provides a pre-defined vocabulary of available verbs (called actions). In addition, a specific event (e.g., assessment event) has a number of pre-defined properties for context that can or must be specified. The properties available vary based on the event type. In the case where there is a need to add a context property not available in the specification, there is a generic extension object, where additional properties can be added (similar to extensions in xAPI). When describing experiences not included in the specification, however, the only option is to use the generic (basic) profile, which supports only generic events (which can use any number of properties), but that can only use the verbs from the pre-defined vocabulary. It seems the use of the basic profile will also pose challenges related to interoperability due to its flexibility in terms of using generic events. Furthermore, while waiting for the addition of new profiles, the pre-defined verbs may not cover the actual needs of users. Since it appears that Caliper is more geared toward the big EdTech companies (Griffiths & Hoel, 2016), this becomes a considerable barrier to adoption for smaller vendors and researchers. For projects that need to support experiences outside of the 14 experiences for which there are profiles, it seems that xAPI would be the better choice after all.

In this case study, we have focused our attention on one specific case, the AVT project. Although we have only examined one project, it is one that really strives to integrate multiple data sources. This case study is one of few studies that has emphasized revealing and addressing challenges and limitations in a data specification, focusing on context descriptions. The research is conducted outside of the safe lab environment, meaning we can identify challenges that would otherwise remain unnoticed. While others have previously identified some of the challenges and limitations of xAPI, this seems to be the first paper that systematically examines challenges and limitations of xAPI context, through involving stakeholders having experienced xAPI in a real-world case, and that gives recommendations on how the context descriptions can be enriched through changes to the context structure and other means in order to improve expressibility.

Conclusion and future work

This paper presents an exploratory case study, taking place in a real-world setting, using the AVT project as a case. The research has aimed to systematically identify challenges and limitations of using a current learning activity data standard (i.e., xAPI) for describing learning context with regard to interoperability and data integration. Subsequently, we have provided recommendations, in summarized form, for the identified challenges and limitations. Our research has identified a lack of clarity in how to describe context data in xAPI regarding interoperability and data integration. The recommendations relate not only to description/modelling of context in xAPI, but also to data typing, validation, and documentation, as all these are essential to enhance the expressibility of xAPI context.

Despite xAPI's potential regarding interoperability, we see a tendency in studies using xAPI that most of them describe only data from one data source. Additionally, in the cases where multiple data sources are actually integrated, few reflect on limitations or challenges concerning data descriptions. In order to scale up LA, particularly when integrating data from multiple sources, it is essential to describe data in a coherent way. Therefore, we strongly encourage others in the LA research community, using xAPI for data integration, to try out the recommended solutions in their own projects. Currently, it is not possible to make use of all recommendations since some will require a change to the xAPI/xAPI profile specifications. Related to the recommendations that can be implemented, we especially highlight the use of xAPI profiles to provide vocabularies and to specify shared data typing and validation rules (through statement templates and patterns).

We acknowledge that there are some limitations with our research. Due to the qualitative approach, where we thematically analyzed data from in-depth interviews with a limited number of participants representing different stakeholder perspectives, the findings are based on our study, although many of the challenges are supported by the literature. Thus, although our results may not be generalizable, they are based on a real-life case involving multiple stakeholders and multiple data sources. Furthermore, the recommendations have not yet been detailed in depth, implemented, and validated.

In future work, we will proceed with the next steps in the methodology, including detailing the recommended solutions that are summarized in this paper, and stakeholder validation of the implementable recommendations through using the xAPI and xAPI profile specifications for data descriptions in two separate projects.

Abbreviations

AVT: Activity Data for Assessment and Adaptation; CAM: Contextual Attention Metadata; IRI: Internationalized Resource Identifier; JSON-LD: JSON for Linking Data; LA: Learning Analytics; LMS: Learning Management System; LOCO: Learning Object Context Ontologies; LRS: Learning Record Store; LCDM: Learning Context Data Model; RCM: Rich Context Model; xAPI: Experience API

Acknowledgements

The authors wish to thank the participants for their valuable contributions in identifying challenges and limitations of xAPI.

Authors' contributions
This paper is a part of the PhD project conducted by JS. BW is her main supervisor, and WC is her co-supervisor. WC has been working closely with JS in planning and has assisted in writing. BW has assisted in planning and writing. All authors read and approved the final manuscript.

Funding
This research is a part of Jeanette Samuelsen's PhD funded by the Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway.

Availability of data and materials
To protect the privacy of the participants, the data cannot be shared.

Declaration

Competing interests
There is no conflict of interests related to this manuscript.

Author details
[1]Centre for the Science of Learning & Technology, University of Bergen, P.O. Box 7807, 5020 Bergen, Norway. [2]Department of Information Science & Media Studies, University of Bergen, P.O. Box 7802, 5020 Bergen, Norway. [3]Oslo Metropolitan University, Oslo, Norway.

Received: 2 October 2020 Accepted: 23 February 2021
Published online: 29 March 2021

References

Advanced Distributed Learning. (2017a). xAPI specification. Retrieved from https://github.com/adlnet/xAPI-Spec
Advanced Distributed Learning. (2017b). xAPI specification - part one: About the experience API. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-About.md#partone
Advanced Distributed Learning. (2017c). xAPI specification - part two: Experience API data. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-Data.md#parttwo
Advanced Distributed Learning. (2017d). xAPI specification - part three: Data processing, validation, and security. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-Communication.md#partthree
Advanced Distributed Learning. (2018a). xAPI profiles specification. Retrieved from https://github.com/adlnet/xapi-profiles
Advanced Distributed Learning. (2018b). xAPI profile specification - part two: xAPI profiles document structure specification. Retrieved from https://github.com/adlnet/xapi-profiles/blob/master/xapi-profiles-structure.md#part-two
Advanced Distributed Learning. (2020a). Anticipating the xAPI Version 2.0 Standard. Retrieved from https://adlnet.gov/news/2020/08/06/Anticipating-the-xAPI-Version-2.0-Standard/
Advanced Distributed Learning. (2020b). xAPI authored profiles. Retrieved from https://github.com/adlnet/xapi-authored-profiles/
Apereo. (2020). Learning Analytics Initiative | Apereo. Retrieved from https://www.apereo.org/communities/learning-analytics-initiative
Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016). Recipe for success: Lessons learnt from using xAPI within the connected learning analytics toolkit. In *Proceedings of the sixth international conference on learning analytics & knowledge*, (pp. 378–382).
Betts, B., & Smith, R. (2019). The learning technology manager's guide to xAPI (Version 2.2). Retrieved from https://learningpool.com/guide-to-xapi/
Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford: Oxford University Press.
Claggett, S. (2018). xAPI Game Demo Example Part 1 [Blog post]. Retrieved from https://gblxapi.org/community-blog-xapi-gbl/10-xapi-demo-example-threedigits
CMI-5. (2020). The cmi5 Project. Retrieved from https://github.com/AICC/CMI-5_Spec_Current
Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing*, *5*(1), 4–7.
European Commission. (2017). New European Interoperability Framework. Retrieved from https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf
Griffiths, D., & Hoel, T. (2016). *Comparing xAPI and Caliper* (Learning Analytics Review, No. 7). Bolton: LACE.
Hruska, M., Long, R., Amburn, C., Kilcullen, T., & Poeppelman, T. (2014). Experience API and team evaluation: Evolving interoperable performance assessment. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*.
IMS Caliper Analytics. (2020). Caliper Analytics | IMS Global Learning Consortium. Retrieved from https://www.imsglobal.org/activity/caliper
IMS Global. (2018). IMS Caliper Specification v1.1. Retrieved from https://www.imsglobal.org/sites/default/files/caliper/v1p1/caliper-spec-v1p1/caliper-spec-v1p1.html
IMS Global. (2020). Members | IMS Global. Retrieved August 20, 2020, from https://site.imsglobal.org/membership/members
International Organization for Standardization. (2011). ISO/IEC 19788-1:2011 Information technology — Learning, education and training — Metadata for learning resources — Part 1: Framework. Retrieved from https://www.iso.org/standard/50772.html
JISC. (2020). Learning records warehouse: technical overview: Integration overview. Retrieved from https://docs.analytics.alpha.jisc.ac.uk/docs/learning-records-warehouse/Technical-Overview:%2D%2DIntegration-Overview
Jovanović, J., Gašević, D., Knight, C., & Richards, G. (2007). Ontologies for effective use of context in e-learning settings. *Journal of Educational Technology & Society*, *10*(3), 47–59.
Keehn, S., & Claggett, S. (2019). Collecting standardized assessment data in games. *Journal of Applied Testing Technology*, *20*(S1), 43–51.

Learning Locker. (2020). Aggregation HTTP interface. Retrieved from https://docs.learninglocker.net/http-aggregation/
Lincke, A. (2020). A computational approach for modelling context across different application domains (Doctoral dissertation, Linnaeus University Press). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-93251
Mazzini, S., & Ricci, F. (2011). EAC-CPF ontology and linked archival data. In *SDA*, (pp. 72–81).
Megliola, M., De Vito, G., Sanguini, R., Wild, F., & Lefrere, P. (2014). Creating awareness of kinaesthetic learning using the Experience API: current practices, emerging challenges, possible solutions. In *CEUR Workshop Proceedings*, (vol. 1238, pp. 11–22).
Miller, B. (2018). Profile Recipes vs. xAPI Profiles [Blog post]. Retrieved from https://xapi.com/blog/profile-recipes-vs-xapi-profiles/
Morlandstø, N. I., Hansen, C. J. S., Wasson, B., & Bull, S. (2019). *Aktivitetsdata for vurdering og tilpasning: Sluttrapport. SLATE Research Report 2019-1*. Bergen: Centre for the Science of Learning & Technology (SLATE) ISBN: 978-82-994238-7-8.
Muslim, A., Chatti, M. A., Mahapatra, T., & Schroeder, U. (2016). A rule-based indicator definition tool for personalized learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge*, (pp. 264–273).
Norwegian Centre for Research Data (2020). NSD - Norwegian Centre for Research Data. Retrieved from https://nsd.no/nsd/english/index.html
NVivo. (2020) Qualitative Data Analysis Software | NVivo. Retrieved from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home
Oates, B. J. (2006). *Researching Information Systems and Computing*. London: SAGE publications.
Papadokostaki, K., Panagiotakis, S., Vassilakis, K., & Malamos, A. (2017). Implementing an adaptive learning system with the use of experience API. In *Interactivity, Game Creation, Design, Learning, and Innovation*, (pp. 393–402). Cham: Springer.
Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0105-4.
Schmitz, H. C., Wolpers, M., Kirschenmann, U., & Niemann, K. (2011). Contextualized attention metadata. In *Human attention in digital environments*, (pp. 186–209).
Siemens, G. (2011). 1st international conference on learning analytics and knowledge. Technology Enhanced Knowledge Research Institute (TEKRI). Retrieved from https://tekri.athabascau.ca/analytics/
Sottilare, R. A., Long, R. A., & Goldberg, B. S. (2017). Enhancing the Experience Application Program Interface (xAPI) to improve domain competency modeling for adaptive instruction. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*, (pp. 265–268).
Standards Norway. (2019). Standards Norway. Retrieved from https://www.standard.no/en/toppvalg/about-us/standards-norway/
Standards Norway. (2020). SN/K 186. Retrieved from https://www.standard.no/standardisering/komiteer/sn/snk-186/
Thüs, H., Chatti, M. A., Brandt, R., & Schroeder, U. (2015). Evolution of interests in the learning context data model. In *Design for Teaching and Learning in a Networked World*, (pp. 479–484). Cham: Springer.
Thüs, H., Chatti, M. A., Yalcin, E., Pallasch, C., Kyryliuk, B., Mageramov, T., & Schroeder, U. (2012). Mobile learning in context. *International Journal of Technology Enhanced Learning*, *4*(5-6), 332–344.
Verborgh, R., & Vander Sande, M. (2020). The Semantic Web identity crisis: in search of the trivialities that never were. *Semantic Web Journal*, *11*(1), 19–27 IOS Press. Retrieved from https://ruben.verborgh.org/articles/the-semantic-web-identity-crisis/.
Vidal, J. C., Rabelo, T., & Lama, M. (2015). Semantic description of the Experience API specification. In *2015 IEEE 15th International Conference on Advanced Learning Technologies*, (pp. 268–269).
Wasson, B., Morlandstø, N. I., & Hansen, C. J. S. (2019). *Summary of SLATE Research Report 2019-1: Activity data for assessment and activity (AVT)*. Bergen: Centre for the Science of Learning & Technology (SLATE). Retrieved from https://bora.uib.no/handle/1956/20187.
Wu, Y., Guo, S., & Zhu, L. (2020). Design and implementation of data collection mechanism for 3D design course based on xAPI standard. *Interactive Learning Environments*, *28*(5), 602–619.
Zapata-Rivera, L. F., & Petrie, M. M. L. (2018). xAPI-based model for tracking on-line laboratory applications. In *2018 IEEE Frontiers in Education Conference (FIE)*, (pp. 1–9).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper 3

Samuelsen, J., Chen, W., & Wasson, B. (submitted). Implementing enriched context descriptions for efficient scaling of Learning Analytics.
Journal of Learning Analytics.

Implementing enriched context descriptions for efficient scaling of Learning Analytics

Jeanette Samuelsen[1,2], Weiqin Chen[1,3], and Barbara Wasson[1,2]

[1] Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway
[2] Department of Information Science & Media Studies, University of Bergen, Norway
[3] Oslo Metropolitan University, Oslo, Norway

Abstract

Learning analytics collects data about learners and their situations, offering the potential of optimizing learning and the situation in which it occurs. To provide a more holistic picture of learners, and more useful analysis, data may be integrated that stem from multiple sources, originating from different contexts. Learning activity data standards, such as xAPI, can support data integration and by extension contribute to scaling up LA. Despite their potential, however, research has shown such standards are not widely used for this purpose, which may be related to challenges and limitations that need to be addressed. In the research detailed in this paper we provide implementation of a number of recommendations related to enhancing xAPI context descriptions and expressibility, including providing a hierarchical and unified representation of context, as part of a technical solution. The technical solution, which is realized through the creation of a new xAPI profile, and that follows the affordances and constraints of xAPI, is adapted to the K-12 school adaptivity use case. The technical solution is evaluated through user testing with technical experts that have experience with xAPI data description in a real-world project, including complementary interviews, which allows participants to explore the solution at a general level and consider different aspects related to its implementation. The usability criteria assessed are usefulness, effectiveness, and satisfaction. The evaluation results indicate that the technical solution meets the criteria for usefulness and effectiveness. Furthermore, the results indicate user satisfaction with the solution at a more general level, related to the recommendations, while also pointing out room for improvement at the implemented level.

Keywords: Learning context, xAPI, Learning analytics, Scalability, Interoperability, Data integration, Learning activity data specification

1. Introduction

In education a number of tools and systems are available to support students and their learning. When students interact with such tools and systems, they may generate trace data that are stored within individual data sources. Technological advances such as ubiquitous Internet and smartphones mean that use of digital learning platforms and related tools are no longer constrained to the physical environment of educational institutions. In addition, publicly available Application Programming Interfaces (API)s and devices with sensor technologies imply that data regarding aspects such as the physical environment and various internal states of a learner can be recorded. Thus, data about and related to learners can be generated from a variety of contexts, where a context can be defined as “any information that can be used to characterize the situation of an entity” (Dey, 2001, p. 5). Different data models have been created that enable the description of learning context (e.g., Jovanović et al., 2007; Schmitz et al., 2011; Thüs et al., 2012; Lincke, 2020), where they vary in aspects such as unit of focus (e.g., learner or learning object), flexibility of data registration, and context categorization (flat or hierarchical) (Samuelsen et al., 2021). Learning analytics (LA) is commonly defined as “the 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” (Siemens, 2011). Related to LA, Bakharia et al. (2016) emphasize that it is essential to add contextual data to learning activity data, i.e., data that describes a learner interacting with a learning object in a learning environment, as a means to improve analysis.

Data integration has been identified as “a key driver of value in analytics” (Cooper, 2014). Integration of data from different contexts, where data exist in disparate data sources, can improve areas such as personalization and adaptive learning (De Meester et al., 2018). Many of the questions that stakeholders wish to use LA to attempt to understand, require the combination of data originating from multiple sources (Kitto et al., 2020). Integration of data is also an important factor of LA scalability (Samuelsen et al., 2019).

Closely related to data integration is interoperability, which can be classified according to semantic, technical, legal, and organizational levels (European Commission, 2017). Within LA there are two well-known educational data specifications (de facto standards), i.e., learning activity data specifications, that address interoperability at the semantic and technical levels, namely Experience API (xAPI; Advanced Distributed Learning, 2021a) and Caliper Analytics (1EdTech, n.d.). Both specifications utilize profiles that enable the definition of vocabularies and enable learning activity data to be governed by specific rules. More information on the specifications and a detailed comparison is available in Griffiths and Hoel (2016).

Despite the promise of these data specifications with regard to data integration, previous research has indicated that they are not widely used for combining data from multiple sources within LA in the context of higher education (Samuelsen et al., 2019). For xAPI, one potential reason for this lack of adoption in terms of data integration may be that the xAPI standard has a number of gaps and needs that can hinder consistent data descriptions across multiple data sources (Samuelsen et al., 2021).

The research identifying xAPI gaps and needs that may be a barrier to consistent data descriptions was the first of a two part exploratory case examining the challenges of using a current learning activity data standard (i.e., xAPI) for describing the learning context. The case is the Activity data for Assessment and Adaptation (AVT) project (Morlandstø et al., 2019; Wasson et al., 2019), which explores how data from multiple EdTech vendors can be integrated and analyzed to provide learner adaptivity. The first part of the case study resulted in recommendations to address the identified gaps and needs, focusing on xAPI context descriptions and expressibility, meaning it should be possible to describe data in a consistent way across data sources.

The research described in this paper stems from phase two of the case study where the following activities were carried out: 1) providing implementation of the recommendations that address the identified gaps and needs regarding xAPI context descriptions and expressibility, as part of a technical solution; and 2) evaluating the usability of the technical solution using technical experts. The search for the technical solution is based on a study of the xAPI and xAPI profile specifications, considering the available functionality and constraints. To accommodate the AVT project case, the technical solution is adapted to the K-12 school adaptivity use case.

2. Background

This section provides information on xAPI that is related to the technical implementation of the recommendations, and the AVT case. Starting with xAPI, the structure of the xAPI statement is explicated, and more information is given regarding xAPI profiles and their functionalities. Subsequently the case is detailed, including how its focus on adaptivity and assessment relates to existing structures of xAPI. Finally, we revisit the recommendations that were implemented as part of the technical solution.

2.1 xAPI

xAPI enables description, integration, and exchange of learning activity data, through the xAPI specification (Advanced Distributed Learning, 2021a) and the xAPI profile specification (Advanced Distributed Learning, 2018a). The xAPI specification provides the general structure to describe learning activity data through xAPI statements, while the xAPI profile specification mainly concerns itself with the content of the statements by addressing the needs of specific use-cases (Advanced Distributed Learning, 2017a).

2.1.1 The xAPI statement

The xAPI statement, described using the JSON format, consists, at the most general level, of the structure *"actor verb object",* meaning that someone (typically a learner) did something, e.g., "learner (*actor*) answered (*verb*) an exam question (*object*)". The object is typically of the type xAPI Activity, which represents something that an actor has interacted with. The interpretation of an Activity is broad, where it can represent both virtual and concrete objects (Advanced Distributed Learning, 2017a). Figure 1 shows the general structure for an xAPI statement (Vidal et al, 2015).

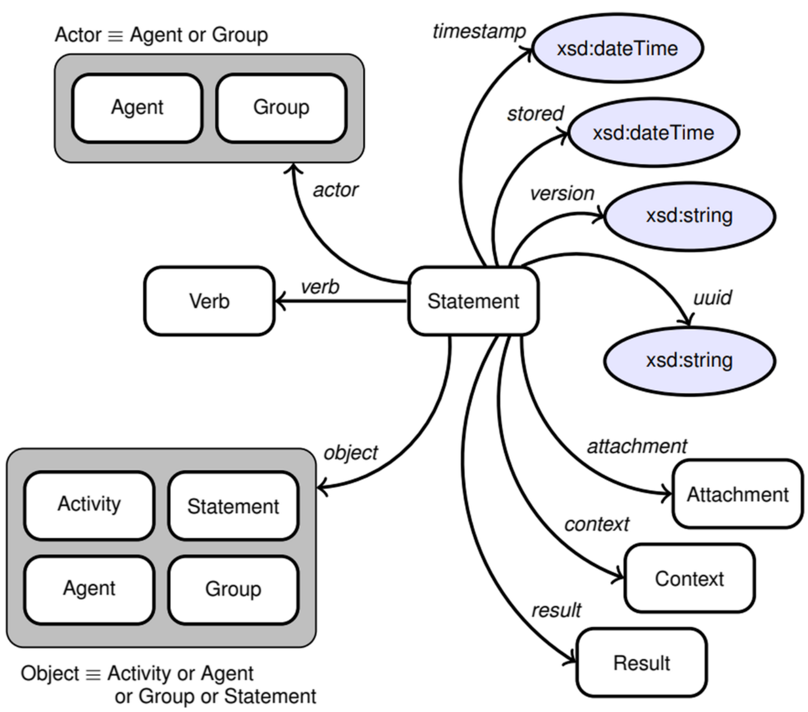

Fig. 1 - xAPI statement (Vidal et al., 2015)

The Activity object includes the required property *id* (an IRI) and *definition* (a JSON object that can contain metadata). The Activity definition object has three recommended properties, *name*, *description*, and *type* (IRI used to define the type of Activity), and includes optional properties, e.g., (Activity) *extensions,* i.e., a JSON object that can contain any properties as needed (Advanced Distributed Learning, 2017b). The number of available properties and sub-properties for the Activity signifies that this object can become long and complex, impacting human-readability and leading to duplication across statements. Thus, xAPI allows for Activities to be stored in the statements themselves, but also includes the option of storing the Activity definition object in the Activity id IRI (referred to as the hosted metadata solution). With the latter solution, the LRS may act as a metadata consumer, thus downloading metadata from the IRI rather than finding it in the xAPI statement (Advanced Distributed Learning, 2017b).

The context structure, which exists on the same level as the actor, verb, and object structures, allows for registration of contextual information related to the learning activity. This structure has 9 properties, where 7 of them (e.g, instructor and platform) accept a single value or single object representing one entity. Context data that are not suitable as values for these 7 properties can be placed in the contextActivities or (context) extensions structures. ContextActivities enables the inclusion of a “map of the types of learning activity context that this Statement is related to” (Advanced Distributed Learning, 2017b). This structure utilizes the xAPI Activity, typically used for the xAPI statement object, to also store contextual information. The Activities can be stored in the learning activity context types: *parent*, *grouping*, *category*, and *other*. The extensions structure enables the addition of a “map of any other domain-specific context relevant to this Statement” (Advanced Distributed Learning, 2017b). The key for an extension must be an IRI, while the value may be any JSON data structure. When using the data structure *JSON object* as the value, it is possible to include multiple levels (properties and sub-properties) of contextual information, with corresponding values (Advanced Distributed Learning, 2018b). The result structure, another structure on the same level as the context structure, allows for registration of context regarding a measured outcome (Advanced Distributed Learning, 2017b).

2.1.2 xAPI profiles

xAPI profiles are specified in the JSON for Linking Data (JSON-LD)[1] format, which builds on JSON and semantic technologies, meaning term definitions can be re-used from established

[1] https://json-ld.org/

vocabularies. For instance, the Simple Knowledge Organization System (SKOS)[2] is re-used in the xAPI profile specification. For individual xAPI profiles, communities of practice may define their own vocabulary concepts, such as verbs, activity types, and extensions. The concepts may be newly created or re-used from other profiles. Profiles can also include statement templates and patterns. Statement templates “provide a set of instructions for what data is required in an xAPI statement which describes a particular event” (Advanced Distributed Learning, n.d.-a). They can include rules that are checked by an xAPI profile validator (human or machine), where the rules may regard presence (inclusion or exclusion) of properties or property values according to statement vocabulary concepts. Patterns define how a group of statements, as defined by statement templates, should be ordered (Advanced Distributed Learning, 2018b). Specific to definition of extensions, the JSON Schema[3] specification can be used within individual xAPI profiles to specify constraints on (sub)properties and values (e.g., value type, [regular expression] pattern and format). Such constraints may be defined in a profile for each extension concept using the *inlineSchema* or *schema* property, and may be more finely specified than what is possible using statement template rules.

2.2 AVT

The research case, the AVT project, is a Norwegian research and development project that originally ran from 2017 to 2019, and subsequently received funding to continue until the end of 2023. The project explores data sharing between K-12 EdTech products for cross-product LA that can be used to provide recommendations for learner adaptivity, and further explores how the data and analysis results may be used for formative assessment. Data coming from multiple EdTech vendor tools are transformed according to a common standard (i.e., xAPI) to enable description, integration, and exchange of data. The types of data to include were informed by research (Morlandstø et al., 2019; Hansen et al., 2020), with interviews with teachers on what they would like to know about their students as input to a learning analytics dashboard being developed in the AVT project. For consistent data descriptions, the project re-used existing xAPI vocabulary concepts and added concepts not previously described into the AVT xAPI profile[4]. The vocabulary concepts, including school and school owner, are defined in the profile as

[2] https://www.w3.org/TR/skos-reference/
[3] https://json-schema.org/
[4] https://github.com/KS-AVT/avt/blob/AVT2/avt.jsonld

activity types. A number of example xAPI statements has also been developed for the project, adhering to the AVT profile (Samuelsen et al., 2021).

The AVT xAPI profile addresses both adaptivity and formative assessment. Formative assessment needs have been addressed in xAPI through the (Activity) *Interaction properties* and the *result* structure[5]. Additionally, some of the published xAPI profiles, e.g., the SCORM profile[6], include concepts related to assessment. Regarding adaptivity, i.e., adapting learning to the individual learner, there are fewer available vocabulary concepts for xAPI, e.g., regarding learner context. Thus, the new vocabulary concepts added to the AVT xAPI profile were mainly related to the use case of adaptivity, although the profile also supports collecting data related to assessment. Further details on the AVT project case can be found in Samuelsen et al. (2021).

The technical solution developed as part of the research reported in this paper, and that implements recommendations that address identified gaps and needs regarding xAPI context descriptions and expressibility, is enabled through the creation of a new xAPI profile, i.e., the school adaptivity profile. The technical solution is customized to specifically address the use case of adaptivity at the K-12 level, and is validated through user testing with technical experts from the AVT project that have experience with xAPI.

2.3 Recommendations

In previous work (Samuelsen et al., 2021), we identified challenges of using xAPI for learning context description through a systematic analysis of data originating from 1) interviews with the stakeholders within the AVT project, and 2) inspection of the xAPI and xAPI profile specifications. Thereafter, we provided recommendations (recommended solutions) to support interoperability and data integration, with emphasis on descriptions of xAPI context, while also providing some recommendations that relate more generally to the xAPI framework (i.e., recommendations for data typing and validation, and documentation). For the list of recommendations, please refer to Samuelsen et al. (2021).

[5] Although there may be semantic differences in how the result structure is used, as mentioned in Samuelsen et al. (2021).
[6] https://profiles.adlnet.gov/profile/313ca4e5-d59e-4b0a-866a-0ad98cc2d1b9

3. Method

This research is based on an exploratory case study methodology (Oates, 2006; Bryman, 2012), where Figure 2 conveys the seven research steps. The first four steps of the case study were addressed in previous research (Samuelsen et al., 2021), while the final three steps form the research that is presented in this paper. Here, a technical implementation of the recommendations, adhering to available functionality and constraints in the xAPI and xAPI profile specifications, was carried out through development of a technical solution. Furthermore, the technical solution was validated, based on user testing that included semi-structured pre-test and post-test interviews.

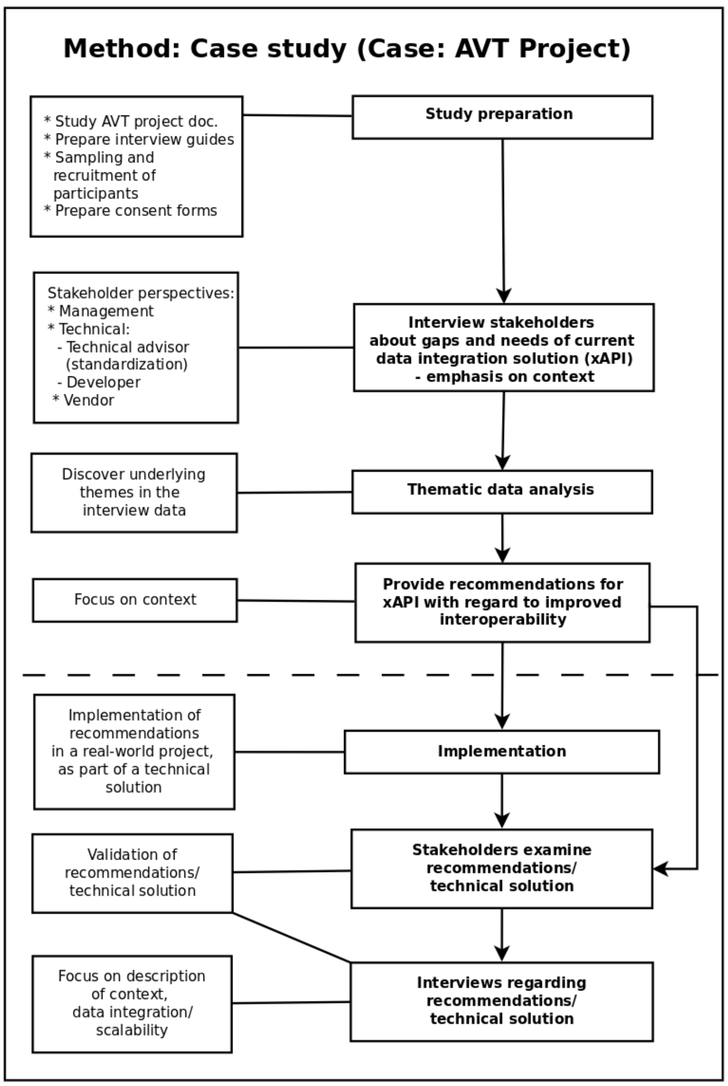

Fig. 2 - Research methodology (adapted from Samuelsen et al., 2021[7])

[7] Fig. 2 has been adapted to reflect that a technical solution implemented the recommendations, and that the technical solution was employed within one real-world project as part of the case study.

4. Technical implementation

In this section the process of providing technical implementation related to the recommendations for interoperability and data integration, through development of a technical solution, is elaborated. Initially, information on how the recommendations relate to the xAPI and xAPI profile specifications more broadly, and on how the recommendations were adapted for implementation according to these specifications, is provided. Next, information on the profile use case, including the context dimensions and properties that were included in the profile, is given. Thereafter, more implementation specific information related to the properties included in the profile is provided, as well as the rationale for 1) the way the properties were implemented according to the affordances of the xAPI profile specification, and 2) two extensions of xAPI constructs and concepts that were added to the technical solution. Finally, the process of automatically validating the profile for conformance with the xAPI profile specification, using tools, is briefly described.

4.1 A brief description of challenges and the technical solution

Based on the recommendations for interoperability and data integration that were at the conceptual level, Table 1 summarizes the technical solution, with aspects grouped according to the challenges identified in Samuelsen et al. (2021). If applicable, aspects of the technical solution are listed both at the general level, relating to the recommendations and how they could apply more generally to the xAPI and xAPI profile specifications, and at the implemented level. In the latter instance, we have explored how the solution requirements can be expressed both using the developed xAPI profile (i.e., the school adaptivity profile) and using xAPI statements that conform to the profile.

Table 1 - Challenges identified and the technical solution

Challenges identified (Samuelsen et al., 2021)	**General technical solution**	**Implemented technical solution**
The distinction between ContextActivities and extensions appears artificial, and it is not always clear which to use.	The xAPI specification should be changed regarding the data model of the xAPI statement context, to allow for context dimensions and their properties rather than ContextActivities with	Use the current xAPI functionality of *context extensions* to illustrate the 2-level hierarchy of context dimensions and their properties. The extension key is the name (represented
Extensions are flexible in how data can be registered and		

therefore could make it more difficult to integrate data.	groupings, and extensions. Through a change to the xAPI profile specification it should consequently be possible to require the inclusion of context dimensions and properties in statements, for the appropriate use cases/community of practice. Following, data typing and validation could be related to the modified xAPI profile, e.g., to ensure the right granularity is achieved in terms of the concepts used.	with an IRI) of the context dimension, and the extension value is a JSON object that contains all the context properties with corresponding values. Since requirements related to extensions within the xAPI profile can be given, data typing and validation are possible, e.g., in terms of deciding if a property should be optional or required, and to ensure the property values follow a certain type or format.
ContextActivities is not a good fit for all types of context data.		
Grouping of related activities in ContextActivities can be done within three different structures (*grouping*, *category* or *other*). It is not clear how the three structures differ. The grouping structures are all very high-level.		
The query capabilities of LRSs are seen as limited. The example given is that it is only possible to filter statements on contextual data that are instances of an activity type. Thus, to allow filtering of resources in AVT (without extra development work), resources were registered as activities, even if they were not really activities on the semantic level.	Query capabilities in xAPI should be altered within the xAPI specification to accommodate the modified context structures (context dimensions with properties solution) that is part of the technical solution.	Not implemented
Concepts from the xAPI vocabulary not sufficient to describe all data.	In addition to the currently available xAPI solution of adding new concepts to xAPI profiles, the xAPI profile specification should support profile inheritance. Enabling a profile hierarchy would allow for having a more generic profile for a use-case/community of practice, which could then be extended by different sub-communities. The more generic profile could have general terms that apply to everyone in the community, while profiles for sub-communities could add specialized vocabulary (e.g.,	Since profile inheritance is not supported, needed concepts (whether new or re-used) are added in the one xAPI profile.

	properties and possible values related to local conditions or practices). The profile inheritance could also apply to data typing and validation, for instance a more general profile only specifies which properties are available for a context dimension, while a more specific profile indicates if certain properties are required, with their data types and value requirements.	
The same vocabulary concept may be represented in different public xAPI profiles, which make up the xAPI vocabulary.	Challenge should be addressed through the profile curation/approval process level rather than a technical solution.	-
Tools generating data at multiple levels of granularity is a challenge, which may make it more difficult to meaningfully integrate data.	For individual profiles we need to be able to specify the type of data values. This functionality is available within the xAPI profiles specification.	For the context extensions used to specify context dimensions, we can specify granularity in a xAPI profile using the JSON schema standard. This adds the possibility of JSON schema validation.
Difficulties in mapping real-world data to xAPI due to its assumptions.	For the general xAPI data model, which is part of the xAPI specification, the assumptions for each property should be made explicit, and tested against the xAPI data through data typing and validation. The xAPI data model should also accommodate new needs as they are identified by the community (Samuelsen et al., 2021).	Related to data typing and validation of the xAPI data model, we use an xAPI statement validator. Concerning accommodating new community needs (general technical solution), currently it is only possible to add additional properties to the xAPI data model on a case-by-case basis for some of the structures (e.g., context, Activity) through extensions. Which extensions should be included in an xAPI statement, and the extension sub-properties, can be specified in an xAPI profile

		using the JSON schema specification (included in the profile through each defined extension's inlineSchema or schema property).
Different tools may generate different numbers of statements for the same type of event.	For individual profiles, the types of data that are required for a statement representing an event and the pattern such statements may follow, can be specified. This functionality is available within the xAPI profiles specification.	xAPI profile statement templates allow us to specify a statement structure.
The openness and flexibility of xAPI allows data and relationships of the same type to be modelled in a myriad of different ways.	Data typing and validation of properties and property values can help alleviate the challenge of openness and flexibility, addressed through a xAPI profile. We can also specify which vocabulary to use through profile vocabularies (Samuelsen et al., 2021).	Use of an xAPI profile to specify vocabulary, data typing and validation.
It may be challenging to correctly use concepts from the xAPI vocabulary.	Challenge should be addressed on documentation level rather than technical solution.	-

4.2 The profile use case

To develop the profile for this research, i.e., the school adaptivity profile, we relied on the guidelines by ADL (Advanced Distributed Learning, n.d.-c). According to these guidelines it is important to identify the use case and requirements for which one would like to utilize xAPI, what should be tracked, and with which concepts this tracking can be achieved. The use case should then inform the profile development. In the following, the profile use case is further outlined.

The profile is informed by the AVT project, focusing on data relevant for K-12 school adaptivity. Currently there is no xAPI profile that explicitly addresses the adaptivity use case. The findings regarding what contextual information teachers would like to know (Morlandstø et al, 2019; Hansen et al., 2020) is important. As the use of context dimensions was one of the main recommendations for data integration and interoperability, a number of context dimensions and properties were identified and classified, primarily derived from the teacher interviews. Additionally, information that could serve as important input to an adaptive system that dynamically changes the learning resource and/or tool offered to the student was taken into account (e.g., the device and environmental context). These two contextual categories are included as learning may happen outside of the school boundaries, and they have been found to be important in previous research (Thüs et al, 2014; Lincke, 2020). Some properties related to the context dimensions were already defined as part of xAPI, for example within the context structure (platform and instructor properties), and as part of profile context extensions published within the profile server (including latitude and longitude properties[8]). Adding such properties to one of the context dimensions enabled explicitly signifying their belonging to a context dimension and their relationships to the other context dimension properties. Finally, contextual information related to the statement itself (i.e., profile and tag properties) and to the object parent was also included. This was done to illustrate how this information could be modeled using the context dimensions and properties approach[9], something that would be needed after the recommended removal of the contextActivities structure.

Table 2 shows the selected context information for the adaptivity profile, modeled as context dimensions with corresponding properties. The unit of measure has been added to a number of the properties to clarify their meaning.

Table 2 - Context dimensions and properties - school adaptivity profile

School/class context	Contains information about the school and class of the statement actor	school, school owner, grade, class, instructor
Platform context	Contains information about the platform, e.g., e-learning platform or LMS, and the relevant structures that might be used to organize content within the platform	platform, feide clientinfo, course, module

[8] These properties are available within the TinCan Vocabulary: https://profiles.adlnet.gov/profile/9596a646-dd44-45fd-92e5-e8cd1a5c6d66.
[9] In the statement context dimension the tag property has the data type "string" rather than Activity.

Statement context	Contains information about the statement, such as the profile the statement is structured according to/follows rules for, and tags (keywords) that pertain to the statement	profile, tag
Environmental context	Contains information related to the external environment of the learner, typically obtained through sensors	latitude, longitude, noise level dB (unit: decibel), linear accelerationX (unit: m/s^2), linear accelerationY (unit: m/s^2), linear accelerationZ (unit: m/s^2)
Device context	Contains information about the device of the learner	connection type, effective connection type, screen width (unit: pixels), screen height (unit: pixels), battery level
Object activity context	Highly dynamic data for the learning object, which relates directly to the learner that is engaging in the activity	parent

Generally, for properties to be added to the profile context dimensions (Table 2), their value should change dynamically relative to the task/item on which the learner works. Properties with values that may be less dynamic in their nature, such as school, are added due to their relevance for reporting within an adaptive system. The properties identified within AVT relate to a Norwegian context, consequently there may be properties listed that would not apply in a similar context in other countries[10]. The object activity context includes only the parent of the learning object (Activity), since the parent of a learning object can change dynamically (e.g., if learning resources are re-used). Static learning object information is not included as this type of information relates directly to the learning object itself (rather than the learner interacting with the learning object), and is addressed through other standards than xAPI (e.g., through IEEE P2881 "Standard for Learning Metadata"[11]).

[10] If the recommended (general) technical solution of profile inheritance was introduced in the xAPI specification, we could make a more generic school adaptivity profile that was extended by local sub-communities such as the AVT project.
[11] https://sagroups.ieee.org/2881/

4.3 Profile implementation

For profile implementation, the information on how the recommendations to support interoperability and data integration can be implemented using the xAPI and xAPI profile specifications (Table 1, column 3) was combined with information regarding the profile use case. Table 3 presents the context properties, grouped according to their respective context dimensions, with information on the specified data types and patterns (where applicable; thus, enabling data typing and validation), activity-types[12] (if applicable), and whether the properties are required. The data types Activity and Agent are from the xAPI specification[13], while the other types (e.g., string, number, array, and object) are from the JSON schema specification.

Table 3 - Context properties/data types (grouped according to context dimensions)

School/class context school (Activity, activity-type: https://w3id.org/xapi/avt/activity-types/school) - required school owner (Activity, activity-type: https://w3id.org/xapi/avt/activity-types/school-owner) - required grade (integer) - required class (string, pattern [regular expression]: [A-Z]) instructor (Agent) - required
Platform context platform (Activity, activity-type: https://w3id.org/xapi/school_adaptivity/activity-types/platform) - required feide-clientinfo (Activity, activity-type: https://w3id.org/xapi/avt/activity-types/feide-clientinfo) course (Activity, activity-type: http://adlnet.gov/expapi/activities/course) module (Activity, activity-type: http://adlnet.gov/expapi/activities/module)
Statement context profile (Activity - activity-type:http://adlnet.gov/expapi/activities/profile) - required tag (array(string))
Environmental context latitude (number, minimum: -90, maximum: 90) longitude (number, minimum: -180, maximum: 180) noise_level_db (number, minimum: 0) linear_accelerationX (number) linear_accelerationY (number) linear_accelerationZ (number)

[12] While IRIs are used to identify the activity-types, the IRIs pertaining to the school adaptivity profile and the AVT profile do not currently resolve, as the profiles have not yet been made publicly available.
[13] The Activity and Agent types were modeled in the profile, as part of the context dimensions, using JSON schema to enable data typing and validation.

Device context connection_type (string, enum: ["bluetooth", "cellular", "ethernet", "mixed", "none", "other", "unknown", "wifi", "wimax"][14]) - required effective_connection_type (string, enum: ["slow-2g", "2g", "3g", and "4g"][15]) battery_level (number, minimum: 0, maximum: 1) screen_width (integer) screen_height (integer)
Object activity context parent (Activity)

The technical solution was implemented using a number of constructs that are defined in the xAPI and xAPI profile specifications. To model context dimensions and accompanying properties and property values in the xAPI statements, xAPI *context extensions* were used. Thus, enabling the expression of a unified structure that accommodates a 2-level hierarchy of context dimensions having properties with values of a variety of possible data types.

Using an xAPI profile allowed us to define the context dimensions to include (by requirement or optionally) in an xAPI statement following the profile and using a specific verb, i.e., *answered*[16], which was demonstrated using *statement templates*. Requirements (or constraints) for individual *context extensions*, including their (sub)properties and values, were defined using JSON schema, with individual schema included using IRIs through the *schema* property. A number of properties were added for the context dimensions, along with information related to their value data types and value patterns.

The xAPI Activity was used as the construct for a number of the context property values within the proposed solution (i.e., when there exists a public IRI that uniquely identifies the value), allowing us to re-use a number of previously defined xAPI Activity types from the AVT profile and other profiles that had defined activity types that were of relevance. Metadata regarding the Activity *definition* was stored in the Activity id IRI, utilizing the hosted metadata solution. Primitive values are used for properties when there is no publicly available IRI to uniquely

[14] Enum values have been gathered from Network Information API (https://wicg.github.io/netinfo/#connectiontype-enum).
[15] Effective connection type represents the measured network performance. Enum values have been gathered from the Network Information API, and while cellular connection types, they apply also for Wi-Fi and other non-cellular connections. An explanation of each value is given in https://developer.mozilla.org/en-US/docs/Glossary/Effective_connection_type.
[16] IRI: http://adlnet.gov/expapi/verbs/answered

identify the resource. Two such properties containing primitive values were grade and class, represented with integer and string, respectively. Since there may be multiple tags related to an xAPI statement, this property takes a JSON array.

To demonstrate how statements could adhere to the profile, we developed an example xAPI statement. The example was partly based on an example made earlier in the AVT project (AVT, 2021); however, the structure and information were modified to conform to the profile developed for this research, i.e., the school adaptivity profile.

4.3.1 Extension of xAPI constructs

For the technical solution, we provided two extensions to the xAPI specification, through emulating functionality. The extensions are related to: 1) the concept of hosted metadata; and, 2) the construct of the xAPI Activity. Both extensions were introduced to enable the consistent use of the hosted metadata solution. Statements using these extensions, with data described according to the functionality that is part of the technical solution, will not be rejected by LRSs. The xAPI specification, however, does not provide semantics to properly understand these extensions.

First, the concept of hosted metadata was extended to also apply for *context extensions,* meaning metadata is hosted in the Activity ids that exist within the context extensions values. Second, the xAPI Activity definition was extended to include data about the metadata IRI that uniquely identifies an object (Activity) within the domain of the administering/controlling organization, using the "authoritative_iri" property. Using such a property could, for example, enable the Activity definition object to be populated according to this information. The property was implemented using (Activity) extensions.

4.4 Automatically validating the xAPI profile using tools

To ensure the technical solution conformed to the xAPI profile specification, two tools were used for profile validation[17]. The first tool is Data Simulator for TLA (DATASIM), which provides functionality to generate simulated xAPI data (Advanced Distributed Learning, n.d.-d), where the xAPI profile that is input to the simulation process is validated (Advanced Distributed Learning,

[17] The reason for using two tools was that they varied in the JSON schema versions used for validation.

n.d.-d). The second tool is the xAPI profile server, providing an endpoint for validating profiles[18] (Advanced Distributed Learning, n.d.-b). To allow for JSON schema validation, the JSON schema that were included using IRIs in the profile were transformed by escaping quotation marks, and included inline in the profile using the individual extensions' *inlineSchema* property. We initially experienced a case where entering newlines within the inline schema was accepted (validated) by DATASIM, but not by the xAPI profile server. After removing newlines from the JSON schemas in the xAPI profile, this issue was resolved.

5. Stakeholder validation

The technical solution was evaluated through user testing, allowing people that are in a product's target audience to review and/or try out the product, and letting them evaluate the extent to which it meets specified usability measures (Rubin & Chisnell, 2008, p. 21).

5.1 Goals

The overall objective of user testing was to assess the usability measures usefulness, effectiveness, and satisfaction for the technical solution. Usefulness relates to the degree a solution allows the users to reach their goals, and assesses willingness to use the product. Effectiveness relates to the degree the product behaves in line with the users' expectation and how simple it is to use the solution, and is often given as a quantitative measure. A part of effectiveness is learnability. Satisfaction relates to how users think about, feel about, and perceive the solution (Rubin & Chisnell, 2008, p. 4-5).

5.2 Participants

Participant selection was conducted according to purposive sampling (Bryman, 2012, p. 418). As the participants needed to understand the xAPI data model, we recruited technical experts who had experience with xAPI data descriptions. Three participants were recruited, having worked within the AVT project as either technical advisors or as a developer for a vendor that contributed data to the project (see Table 4).

[18] https://profiles.adlnet.gov/api-info/validate

Table 4 - Participants

Identifier	Gender	Perspective (AVT)	Work tasks in AVT (sample)
P1	Male	Technical advisor	Work related to vocabularies and the AVT xAPI profile, examined how to represent context for xAPI activity data, development of xAPI example statements for use by the vendors
P2	Male	Technical advisor	Work related to vocabularies and the AVT xAPI profile, examined how to represent context for xAPI activity data, participated in developing the AVT example xAPI statements, advised on tools and libraries for xAPI activity data storage and exchange
P3	Male	Vendor, developer	Technical work on behalf of the vendor, including implementing the vendor's solution for transformation of activity data to xAPI, delivery of xAPI statements to the AVT project

The participants that had functioned as technical advisors in the AVT project (P1, P2) had been involved in the first phase of the project (2017-2019). P3 was involved in data delivery to the project in 2022.

A key characteristic of the participant group is that they are users with highly specialized technical skills and extensive experience with xAPI, and could therefore be expected to go through the user testing based on materials (e.g., test instructions, documentation) without guidance from the test moderator.

5.3 Study design

Materials were developed and piloted prior to the user testing. The materials included a data registration form that among other described user test tasks that would result in an xAPI statement that would be structured according to the school adaptivity profile, and pre-test and post-test interview guides.

A user test was designed comprising 1) a pre-test interview, 2) a testing session, and 3) a post-test interview. The user test was conducted to gain an indication of the technical solution's usability, according to the usability goals. The user test was designed as an assessment test, which enables evaluation of the usability of a solution's lower-level functionality and aspects, and uncovering usability deficiencies, through users performing realistic tasks (Rubin & Chisnell, 2008, p. 34-35).

5.3.1 Ethical approval

The project was reported to the Norwegian Agency for Shared Services in Education and Research (Sikt) - Data Protection Services[19], which assesses that the processing of personal information is legal based on the measures taken concerning privacy and research ethics. Participant consent was obtained through a signed consent form to collect data from audio recorded interviews and for their participation in the user testing tasks.

5.4 Carrying out the user testing

The user testing was conducted between November 23 and December 16, 2022. Data were collected from 1) pre-test interviews, covering qualitative data regarding participant background; 2) testing session, i.e., quantitative performance data (behavioral measures) derived from user test tasks, to identify errors in task completion; 3) post-test interviews, covering qualitative data about the test experience and preference data regarding various aspects of the recommendations that serve as a foundation for the technical solution and its implementation. The interviews, which took place using a video conference platform, were conducted in Norwegian.

5.4.1 Pre-test interviews

Each user test session started with a pre-test interview activity. This was supported by the pre-test interview guide that contained questions about the background of the participants (see section 5.2). The activity concluded with us giving the participants information about the technical solution that would be user tested and what the user testing would entail.

Upon interview completion, the participants were emailed the data registration form and either an anonymized xAPI statement obtained from the AVT project (in case of the technical advisors) or an xAPI statement they had previously delivered to the AVT project (in case of the vendor); each participant received a different statement.

[19] https://www.nsd.no/en/data-protection-services

5.4.2 Testing sessions

The participants were asked to go through the user testing themselves, allowing them to do the test over several days, in their own environment, on their own equipment. The data registration form, detailing the user test, described 6 mandatory and 1 optional data description tasks, which would result in a new xAPI statement. The first task involved filling in the statement's actor, verb, object and result structures, ensuring that the participants had a general understanding of xAPI, while the remaining tasks were focused on specifying context dimensions, properties, and values in accordance with the required value type/format. Additionally, the data registration form included materials that could support the execution of the user test tasks, including documentation regarding the technical solution (e.g., from Table 2 and Table 3) and an annotated version of the xAPI profile, a link to an xAPI statement validator[20], a section for note taking during the tasks, e.g., regarding thoughts and opinions of the solution, and a section where participants could gradually fill in the xAPI statement that would result from completing the tasks.

Participants were asked to base their xAPI statement on the activity data they had been emailed. Structuring the data in accordance with the school adaptivity profile required changes to the data (e.g., in terms of structure and data types), and some additional data needed to be included in the xAPI statement according to fictitious values that would fit the required data value type/format. The participants were encouraged to complete the tasks within 7 days, and to email the completed xAPI statement. Upon receiving the email, we evaluated the user test and set up the post-test interview.

5.4.3 Post-test interviews

The post-test interview was supported by the post-test interview guide, which included relevant topics. The interview was related to the participant's user test, where participants were encouraged to bring their notes, allowing for an improved understanding of the participants' actions and any potential errors or issues. Furthermore, the interview contained questions about different facets related to the technical solution on a more general level, and its implementation, allowing us to ascertain the strengths and weaknesses of the technical solution. These facets included the structuring of the contextual data (using a unified and hierarchical structure), the chosen context dimensions and properties, requirements for context dimensions and properties,

[20] The data registration form linked to the xAPI statement validator available at https://lrs.io/ui/tools/xapi-statement-validator/.

and the hosted metadata solution. Finally, the interview inquired about profile inheritance, to assess if such functionality could potentially be valuable.

5.5 Data analysis and results

Data were analyzed relating to the different parts of the user test, which involved assessing the completion of the required user test tasks, according to quantitative measures. For each of the six required tasks, 1 point was the maximum number of points that could be obtained, leading to a maximum score of 6 for the completed user test. Each task had a number of subtasks, where every task except for the first involved filling out the context dimension with a number of adhering properties. For each task, we checked if the subtasks were completed according to the expectation. Correct completion of a subtask would yield 1 point, where the number of points were summed up for each task and divided by the number of subtasks, giving the task score. The post-test interviews were transcribed, and the transcripts were analyzed qualitatively according to topics, to identify the participants' thinking related to different facets regarding the more general technical solution and its implementation. Quotations, originating from the post-test interview, have been translated from Norwegian to English.

5.5.1 Task completion

A summary of the user test results for the first 6 tasks is shown in Table 5. P2 and P3 completed the optional seventh task correctly.

Table 5 - User test results

Test #	**Participant #**	**Tasks completed**	**Score (out of 6 possible points)**
Test 1	P1	6	5.83
Test 2	P2	6	5.83
Test 3	P3	6	5.33

5.5.2 Validation of the technical solution

Information regarding the users' test experiences and their thoughts and opinions regarding the recommendations that serve as a foundation for the technical solution, and its implementation, derived from the post-test interview and further supported by participants' notes, are provided below. The findings are listed according to topics, where the 7 first topics stem from the post-test interview guide, while the final 2 topics emerged from the post-test interviews.

User test experience

The user test instructions were generally understood by the participants, who found the tasks quite straightforward to go through supported by a number of materials, leading them to create and submit an xAPI statement resulting from the tasks. In the one case where all the participants had chosen a different way to describe the data than what had been intended by the researchers, there was a discrepancy between the user test instructions and the constraints given as part of the technical solution (the relevant user test task asked that participants fill in a property that was not required to be filled in as part of the technical solution). As such, the user test instructions could have stated more clearly that in the case where such differences are observed, the participants should follow the user test instructions. Additionally, one participant (P3) did not correctly fill in the profile that the statement was structured according to, which may be related to him being less familiar with the concept of an xAPI profile than the other participants. Finally, to further support the data description tasks, one participant (P2) suggested statement validation according to the profile and including the xAPI profile JSON schemas explicitly in the support materials.

Structuring of contextual data

All participants expressed approval for the use of a hierarchical structure to classify context properties in accordance with context dimensions. The participants indicated that this structure, which is more explicit than the flat structure typically used in xAPI, adds to aspects such as understanding, readability, and cleaner organization of the properties, thereby promoting expressibility. In terms of the unified structure, the participants also generally agreed that it is preferable to spreading data across structures, mentioning aspects such as readability, understandability, and learnability, indicating that the unified approach could further support expressibility. At the same time, two of the participants (P2 and P3) expressed pragmatic views regarding data integration, suggesting that as an alternative to achieving integration through the unified and hierarchical solution, the different parties may alternatively follow the same rules for consistent data descriptions, or that the xAPI statements may be further processed to achieve integration.

Context dimensions and properties

The participants indicated that 4 of the 6 context dimensions seemed appropriate to include based on the profile's focus on adaptivity in the K-12 domain. For the remaining 2 context dimensions, i.e., the "device" context and the "environmental" context, two of the participants (P1 and P3) indicated they were less certain regarding their inclusion. Here, P3 understood adaptivity in a somewhat different way than the broader understanding that had been the basis for the school adaptivity profile. While the school adaptivity profile included factors related to the learning environment, including device and environmental context dimensions, P3's understanding of adaptivity was that “if the student does not manage to successfully solve a task then he will get an easier task, and if he is able to successfully solve a task then he will get a more challenging task”. Furthermore, P3 also highlighted that universal design of learning resources might reduce the need to adapt to these context dimensions. In terms of adding further context dimensions, P2 indicated that there may be other context dimensions that could support the goal of adaptivity, although he was uncertain of which. In general, P1 and P2 advised that the context dimensions that introduced new information (all except the statement and object activity context dimension) should not be a part of the standard, but should rather be included as needed for particular communities of practice.

The participants suggested possible improvements for the context dimensions:

- The platform dimension needs a better name, as it encompasses the properties of platform, feide-client-info, course, and module, and to avoid duplication of terms across context dimensions and properties (P1 and P2)
- The "parent" of the "object_activity" context might belong together with course and module in the "platform" context, but it would depend on the circumstance (P3)
- The statement context dimension could be renamed to clarify its meaning and how its properties are related (P3)

The participants also suggested possible improvements for the properties - related to context dimensions:

- Choose a more international name for feide-clientinfo to adapt to the international K-12 adaptivity community (P2)
- Find a more fitting measure for movement than linear acceleration (P3)
- The profile property value (IRI) should explicitly clarify the profile version (P1)

- Add properties course "unit" and "deep link" used to navigate to a resource to the platform dimension (P2)
- Add properties "touch screen" (P2 and P3), and "user-agent" and "viewport" (P3) to the device dimension
- Add information regarding universal design and individual adaptation related to disability and special needs (P1)
- The grade property is ambiguous in its meaning and needs clarification (P1 and P3)

Requirements for context dimensions and properties

Regarding the data typing and validation measures, the participants indicated on one hand that such measures may promote interoperability and data integration, but on the other hand they indicated that it is important that the measures must fit reality. The participants identified some properties, including "class" and "instructor", where the data types and format requirements might not fit exactly with reality. One participant (P2) cautioned also that customizing the solution too much for the Norwegian case might make the solution unusable by the international community. Another issue (brought up by P1) was that when creating the technical solution, it was unnecessary to reuse old data types from xAPI (e.g., the Activity data type). At the minimum, he indicated, it should be considered if the Activity could have a more generic name to signal its broader application area.

Further improvement suggestions regarding data types and formats:

- Tag property, which takes a list of string, needs to be standardized (P2 and P3)
- Allow for specifying more than one profile (P2)
- Numeric data types could have further restrictions, e.g., regarding decimal precision and ranges (P2)

There was also indication by the participants (P2 and P3) that when requiring context dimensions and properties, there may be a trade-off between interoperability (through consistent data descriptions) and adapting to real-world situations. P3: “It's the usual challenge, the more things that are required, the more interoperability. [...] But then it gets more difficult to accommodate the real-world”. Another issue mentioned (by P1) is that requiring a lot of data to be present in all xAPI statements of a certain type would impact storage space and data transfer latency.

Further improvement suggestions regarding context dimensions and property requirements:

- It is better not to require environment and object activity context dimensions since their properties are not required (P2 and P3)
- The instructor property might not need to be required (P3)

Implementation of technical solution

The participants commented on a number of functionality within xAPI that were utilized for implementation of the recommendations as part of the technical solution. Statement templates were used to require context dimensions. Some aspects mentioned by the participants included further use of such templates to also require statement properties (e.g., statement id), as an alternative to JSON schema definitions in specifying context dimension properties (as mentioned by P2), use of statement templates with profile patterns to reduce the length of statements within such a pattern (i.e., group of xAPI statements; when applicable, as mentioned by P1), and that while the statement templates increase human-readability there exist very few tools that may actually provide template validation for statements that are structured in accordance to a specific profile (P2).

The xAPI profile utilized xAPI extensions to represent the context dimensions. Within the extensions, individual JSON schemas specified the requirements for each of the context dimensions. One aspect mentioned (by P2) is that it could be stated more explicitly in the context extension IRIs that they are actually context extensions. The participants also had some varying inputs regarding the JSON schema. For instance, one participant (P1) thought it was an interesting way to extend the capabilities of xAPI, but noted that other schema languages (e.g., XML Schema Definition [XSD]) could be more expressive. Another participant (P2) stated that the JSON schema could enforce the use of JSON-LD compatible syntax, if the xAPI statements in the technical solution had been specified using JSON-LD rather than JSON. Regarding including JSON schema in the profile through external IRIs with the "schema" property versus in the profile with the "inlineSchema" property, the participants had a preference to utilize external schemas, listing a number of reasons related to expressibility. At the same time, it was mentioned (by P2) that validation tools may provide better support when the schemas are included with "inlineSchema".

Hosted metadata

The participants indicated approval for the hosted metadata solution, provided a stable infrastructure was used to serve the metadata and a versioning strategy was put in place, with two participants (P2 and P3) signaling they would generally prefer this solution over inlining data about Activities in the xAPI statements.

Some of the benefits mentioned regarding this solution include statement length and consistent data descriptions (P2 and P3), and that it could make it easier to correct mistakes and filter the data (P3). Although the xAPI specification hints at the possibility of deriving metadata from other formats, it does not provide implementation details for such a solution. As such, one participant (P1) alluded to the problem of potential data duplication across standards, and one participant (P2) suggested several approaches to hosted metadata that could be used with more than one standard (e.g., regarding learning analytics and learning resources), including using JSON-LD metadata. As a potential drawback, P2 cautioned that it might be difficult to understand the solution by those that would host the metadata. P2: “It might be hard enough for them to understand xAPI, so having hosted metadata [...] in the object IDs that they thought they would use, that would likely be a challenge”. The participants perceived the "authoritative_iri" property as potentially useful for hosted metadata, listing uses such as data provenance (P3) and retrieving updated metadata (P2).

Profile inheritance

The participants generally expressed that profile inheritance, where one profile through inheritance can add all that is contained in another profile and then add its own specific properties, could be advantageous (e.g., in terms of catering to international communities and extending the applicability of an xAPI profile). However, the participants also highlighted different nuances related to inheritance, and different aspects that would need to be taken into account when considering such an approach. One aspect mentioned (by P1) is that even though one may not directly inherit/extend a profile, it is currently possible to reuse some of the xAPI concepts from another profile. Furthermore, it is possible through the xAPI profile specification to specify how concepts such as verbs and activity types relate to other concepts, since xAPI profiles inherit linked data concepts from the SKOS specification (as emphasized by P2). Additionally, P2 highlighted that inheritance is different in JSON-LD than in object-oriented programming languages, having implications for how the concepts may be inherited/extended in practice.

Documentation

Aspects related to the documentation of the technical solution were brought up during the interviews. At the level of the xAPI profile, both P1 and P2 indicated that more documentation could have been provided regarding the definitions and use of the context dimensions. As expressed by P1: "I was wondering if in the profile it would be possible to insert the definitions from the tables [table 2 and table 3]. It would have been useful." In terms of documenting the semantics for the context dimension properties, P1 and P2 stated that JSON schemas could provide definitions for the individual schema properties. Additionally, P1 indicated that it would be good to include examples of property usage through the JSON schemas.

Linked data

Both P1 and P2 highlighted inclusion of linked data technologies as a potential improvement for the xAPI specification and also for the technical solution. P1 and P2 suggested that it would be good to add a JSON-LD @context to the xAPI statements, which could uniquely identify the JSON-LD properties and clarify their meaning through online lookup, even though the xAPI specification currently uses JSON rather than JSON-LD. Similarly, P1 and P2 stated, using linked data in statements would also allow the learning resources or activities to be referred to uniquely as linked data, providing additional context related to their meaning. Both P1 and P2 presented hosting metadata as linked data as an alternative approach to xAPI hosted metadata, P2 clarifying that the xAPI hosted metadata is limited (in terms of format and use case). In addition, P1 highlighted that JSON-LD is an RDF format, which means it can be imported with ontologies into a triple store, leading to more powerful and intelligent search capabilities than what will be available through the LRS.

6. Discussion

Returning to the user test goals, we can now get an initial indication of the usability criteria of usefulness, effectiveness, and satisfaction. Regarding usefulness, the user test indicates that the technical solution allows the users to reach their goals (of describing contextual data with xAPI in a consistent manner across data sources), where the participants express that the use of context dimensions and properties, as part of an unified and hierarchical solution, add to aspects such as readability and understandability. Two participants express, however, that such usefulness can also be achieved through the existing xAPI and xAPI specifications if care is taken for data to be combined within and across sources (e.g., through following examples/templates for data description). Additionally, the participants indicate support for data typing and validation measures in general, but caution that it is important that the data types and validation measures correspond to the real-world. Next, it is indicated that the effectiveness criteria is met, in that users were for the most part able to correctly understand how to utilize the technical solution (as indicated from the results on the user test), and also with participants expressing that based on the instructions and support materials it was mostly straightforward to go through the data description tasks. Finally, the participants indicated positive attitudes toward the technical solution at a general level, indicating satisfaction with recommendations related to interoperability, data integration, and xAPI context descriptions. The participants, however, identified some room for improvements on the implemented level, where some of these improvements relate only to the technical solution (e.g., documentation and property requirements), while others relate also to the xAPI specification, and its affordances, more generally.

6.1 The use of the xAPI Activity

Related to the use of the xAPI Activity, whose semantics has been found to be problematic by some of the research participants, we advise considering if the name (Activity) could be made more generic, and furthermore, if its definition could be broadened and documented within the xAPI documentation. This would make it clear that this structure may be used for values not just related to the object of an xAPI statement, but also within the context of the statement.

6.2 Validation and LRS compatibility

Some challenges were encountered for the profile validation, including different behaviors among the two tools that were used, i.e., DATASIM and the xAPI profile server. To foster mainstream adoption of xAPI, improvements could be made to the validation tools and capabilities offered. For instance, there needs to be mechanisms to ensure validators that validate according to the same parameters do not produce different results. For the latter aspect, inspiration could be gathered from a previous project that examined statement validation across LRSs, as a means to investigate practical interoperability (Downes et al., 2015).

Furthermore, related to the LRSs' that implement the xAPI specification, xAPI operates with three different conformance levels, i.e., MUST, SHOULD, and MAY (Advanced Distributed Learning, 2017a). In particular, the MAY level could lead to inconsistent behaviors among LRSs, indicating an option where the developer is free to choose without any recommendation from the specification. One such example that we encountered is for the hosted metadata solution, where it is optional for the LRS to act as a metadata consumer that resolves the hosted metadata IRIs (Advanced Distributed Learning, 2017b). We tried this solution with the Learning Locker Open Source LRS[21], and found that the hosted metadata were not resolved. If the specification is to be implemented in an interoperable manner, where data may be integrated across multiple sources, it would be beneficial for xAPI to specify clearer guidelines.

6.3 Extensibility and broader applicability

Aspects related to the extensibility and broader applicability of the solution were discussed during the post-test interview. Two of the participants brought up the need to separate between properties that are included in the standard and properties that are included for specific communities of practice. While we agree with the suggestion that it would not scale well to include all properties in the standard, it might be possible to include the more custom properties as part of context dimensions in profiles, while the more general properties could be made part of context dimensions in the standard. The separation between the standards and profiles could be further explored by a number of relevant stakeholders. Here, the Caliper specification (1EdTech, 2020) could serve as inspiration by looking at its separation between properties that are required for all Events[22] and properties that pertain to specific Caliper profiles.

Another aspect expressed was that some of the property names could be made more generic or internationalized, to accommodate a broader adoption of profiles. Finally, the participants were generally positive to the use of profile inheritance, where one base profile could be inherited by relevant sub-communities (e.g., to accommodate cultural or national subtleties). While the participants indicated that such a solution could include a base profile that is inherited, meaning it is contained in the inheriting profile, further aspects of such a solution would need to be worked out, to assess if the suggestion is feasible both on a conceptual and an implementable level.

[21] https://github.com/LearningLocker/learninglocker

[22] “A Caliper Event is a generic type that describes the relationship established between an actor and an object, formed as a result of a purposeful action undertaken by the actor at a particular moment in time and within a given learning context” (1EdTech, 2020).

6.4 The use of JSON-LD

The participants discussed the potential benefits of using JSON-LD for xAPI statements in the technical solution, and more generally for the xAPI specification, as this approach would allow the use of linked data technologies. Concerning the unified and hierarchical structure that has received positive feedback in the user test, it would still be possible to define it using JSON-LD since JSON-LD relies on JSON which allows for nested data structures at multiple levels. Currently it is only the xAPI profile specification, not the xAPI specification, that supports JSON-LD. While there exists an xAPI ontology, the xAPI community reports that there are major issues with the ontology[23]. Various research has been carried out related to xAPI ontology development (e.g., Vidal et al., 2015; De Nies et al., 2015; Vidal et al., 2018), although it does not appear that results from these projects have been adopted explicitly into the xAPI ontology. If moving to a linked data solution for the xAPI specification, the xAPI community should look into this previous research. Utilizing JSON-LD for the xAPI specification could have a number of possible advantages, which were emphasized by the participants, including enhanced query capabilities in triple stores, hosting metadata that may pertain to different standards, and reducing xAPI statement length, thus potentially increasing expressibility through more consistent descriptions. For linked data, data typing and validation constraints would typically be defined in languages such as Shapes Constraint Language (SHACL), and RDF schema (RDFS), allowing for the use of RDF built-in data types, which are generally more expressive than JSON schema (Allemang et al., 2020). On the other hand, the JSON schema specification, utilized for the technical solution, could also play a role in ensuring xAPI statements are according to the requirements, especially for developers that are familiar with JSON but not with linked data.

7. Conclusion and future work

In this paper we detail the second part of an exploratory case study, using the AVT project as a real-world case. While the first part identified challenges related to xAPI context, interoperability, and data integration, and provided recommendations for improvements to xAPI, emphasizing expressibility, the current research has aimed to provide implementation of the recommendations, as part of a technical solution, whereafter the technical solution has been evaluated by technical experts.

[23] https://github.com/adlnet/xapi-ontology/issues

The technical solution has been implemented according to the functionality and affordances of the xAPI and xAPI profile specifications, through the creation of a new xAPI profile for the K-12 adaptivity domain with an accompanying xAPI example, allowing participants to explore and reason about its different aspects, which includes a unified and hierarchical context model that is exemplified with categorization of context according to the adaptivity domain, and that emphasizes measures for data typing and validation.

The technical solution was validated through user testing, with data description tasks and complementary interviews. The participants generally indicated approval of the unified and hierarchical structuring of context, as represented through context dimensions and properties, emphasizing that it can add to aspects related to expressibility. The participants, however, indicated pragmatic views that data integration can alternatively be achieved through other means such as post hoc data processing. Furthermore, the participants indicated general support for data typing and validation measures, although cautioning that these need to follow real-world constraints. The participants also indicated approval for the use of hosted metadata. Related to the usability criteria assessed, the user testing generally indicates that the usefulness and effectiveness criteria are met. Additionally, the participants indicate satisfaction regarding the solution at a general level, related to the recommendations, while also pointing out potential for improvement at the implementable level.

Expressibility may promote interoperability and data integration, which both relate to LA scalability. As such, measures utilized in the technical solution (e.g., regarding the hierarchical and unified structuring of contextual information), could be further explored for inclusion to the standard by the xAPI community, as a means to promote scalable LA, thereby breaking down the silos that act as a barrier to LA without having to rely on time and resource-intensive ad hoc or post hoc processing of the data. Furthermore, the stakeholder validation results highlight the potential of including JSON-LD for both the technical solution and the xAPI specification (e.g., to provide additional semantics to properties and as a cross-standard compatible approach to host metadata).

We acknowledge there are limitations to this research. While focusing on implementing the recommendations related to interoperability, data integration, and xAPI context descriptions as part of the technical solution, we did not address and implement all of the recommendations, due to the affordances and constraints provided by the current xAPI and xAPI profile specifications, and the research situation. For instance, we do not directly address the need for enhanced query capabilities in xAPI; and the solution does not include xAPI profile patterns, used to specify the ordering of allowed xAPI statements, since the use case of school adaptivity is very broad, meaning a lot of different sequences could potentially be allowed. Furthermore, the usability of the technical solution has been evaluated using technical experts from only one case, the AVT project, with a limited number of participants. Finally, the device and environment context dimensions, while having foundations in previous research, have not been evaluated by domain experts within education, such as teachers. Similarly, the categorization of context properties into context dimensions, and the data types and validation measures, have not yet been validated by educational domain stakeholders.

As future work, the technical solution should be evaluated by a greater number of technical experts, and the context data model should be evaluated by a broader group of stakeholders, having domain expertise within education and educational technology.

References

1EdTech (n.d.). Caliper Analytics. Retrieved October 13, 2022, from https://www.imsglobal.org/activity/caliper

1EdTech (2020). Caliper Analytics® Specification. Retrieved from https://www.imsglobal.org/spec/caliper/v1p2

Advanced Distributed Learning. (n.d.-a). xAPI Profile Server Authoring Guide - Authoring Guidance. Retrieved from https://adlnet.gov/guides/xapi-profile-server/authoring-guide/Authoring-Guidance.html

Advanced Distributed Learning. (n.d.-b). xAPI Profile Server. Retrieved from https://profiles.adlnet.gov/profiles

Advanced Distributed Learning. (n.d.-c). xAPI Profile Server - xAPI PROFILE GUIDELINES. Retrieved from https://profiles.adlnet.gov/help#resourcestop

Advanced Distributed Learning. (n.d.-d). Data Simulator for TLA (DATASIM). Retrieved from https://adlnet.gov/projects/datasim/

Advanced Distributed Learning. (2017a). xAPI Specification - Part One: About the Experience API. Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-About.md#partone

Advanced Distributed Learning. (2017b). xAPI Specification - Part Two: Experience API Data.

Retrieved from https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-Data.md#parttwo

Advanced Distributed Learning. (2018a). xAPI profiles specification. Retrieved from https://github.com/adlnet/xapi-profiles

Advanced Distributed Learning. (2018b). xAPI Profile Specification - Part Two: xAPI Profiles Document Structure Specification. Retrieved from https://github.com/adlnet/xapi-profiles/blob/master/xapi-profiles-structure.md#part-two

Advanced Distributed Learning. (2021a). xAPI specification. Retrieved from https://github.com/adlnet/xAPI-Spec

Allemang, D., Hendler, J., & Gandon, F. (2020). *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL* (3rd ed.). ACM. https://doi.org/10.1145/3382097

AVT (2021). Eksempler på kall mot og respons fra leverandørenes aktivitetsdata-API. Retrieved from https://github.com/KS-AVT/avt/blob/AVT1/eksempler.md

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016, April). Recipe for success: lessons learnt from using xAPI within the connected learning analytics toolkit. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 378-382).

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford: Oxford University Press.

Cooper, A. (2014). Learning analytics interoperability: The big picture in brief. Learning Analytics Community Exchange, March 26, 2014. http://www.laceproject.eu/blog/learning-analytics-interoperability-briefing

De Meester, B., Lieber, S., Dimou, A., & Verborgh, R. (2018). Interoperable user tracking logs using {linked data} for improved learning analytics. In *Call Your DATA Proceedings of the XIXth international CALL research conference* (pp. 29-31).

De Nies, T., Salliau, F., Verborgh, R., Mannens, E., & Van de Walle, R. (2015). TinCan2PROV: exposing interoperable provenance of learning processes through experience API logs. In *Proceedings of the 24th international conference on World Wide Web* (pp. 689-694).

Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, *5*(1), 4-7.

Downes, A., Shahrazad, A., & Smith, R. (2015). *Sharing between LRSs: A collaborative experiment in practical interoperability* [White paper]. https://xapi.com/wp-content/uploads/sites/3/2015/03/whitepaper.pdf

European Commission. (2017). New European Interoperability Framework. Retrieved from https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf

Griffiths, D., & Hoel, T. (2016). *Comparing xAPI and Caliper* (Learning Analytics Review, No. 7). Bolton: LACE.

Hansen, C.J.S., Wasson, B., & Belokrys, G. (2020). *Teacher's need for information.* Internal report: unpublished.

Jovanović, J., Gašević, D., Knight, C., & Richards, G. (2007). Ontologies for effective use of context in e-learning settings. *Journal of Educational Technology & Society*, *10*(3), 47–59.

Kitto, K., Whitmer, J., Silvers, A. E., & Webb, M. (2020). Creating Data for Learning Analytics Ecosystems. *Society for Learning Analytics Research (SoLAR)*. Retrieved from https://www.solaresearch.org/wp-content/uploads/2020/09/SoLAR_Position-Paper_2020_09.pdf

Lincke, A. (2020). *A Computational Approach for Modelling Context across Different Application Domains* (Doctoral dissertation, Linnaeus University Press). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-93251

Morlandstø, N.I., Hansen, C.J.S., Wasson, B., & Bull, S. (2019). Aktivitetsdata for vurdering og tilpasning: Sluttrapport. SLATE Research Report 2019-1, Bergen, Norway: Centre for the Science of Learning & Technology (SLATE). ISBN: 978-82-994238-7-8

Oates, B. J. (2006). *Researching Information Systems and Computing.* SAGE publications.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons.

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0105-4

Samuelsen, J., Chen, W., & Wasson, B. (2021). Enriching context descriptions for enhanced LA scalability: a case study. *Research and Practice in Technology Enhanced Learning, 16*(1). https://doi.org/10.1186/s41039-021-00150-2

Schmitz, H. C., Wolpers, M., Kirschenmann, U., & Niemann, K. (2011). Contextualized attention metadata. In *Human attention in digital environments* (pp. 186–209).

Siemens, G. (2011). 1st international conference on learning analytics and knowledge. Technology Enhanced Knowledge Research Institute (TEKRI). Retrieved from https://tekri.athabascau.ca/analytics/

Thüs, H., Chatti, M. A., Yalcin, E., Pallasch, C., Kyryliuk, B., Mageramov, T., & Schroeder, U. (2012). Mobile learning in context. *International Journal of Technology Enhanced*

Learning, *4*(5-6), 332–344.

Thüs, H., Chatti, M. A., Greven, C., & Schroeder, U. (2014). Kontexterfassung,-modellierung und-auswertung in Lernumgebungen. *DeLFI 2014-Die 12. e-Learning Fachtagung Informatik*.

Vidal, J. C., Rabelo, T., & Lama, M. (2015). Semantic description of the Experience API specification. In *2015 IEEE 15th International Conference on Advanced Learning Technologies*, (pp. 268–269).

Vidal, J. C., Rabelo, T., Lama, M., & Amorim, R. (2018). Ontology-based approach for the validation and conformance testing of xAPI events. *Knowledge-Based Systems*, *155*, 22-34.

Wasson, B., Morlandstø, N. I., & Hansen, C. J. S. (2019). *Summary of SLATE Research Report 2019-1: Activity data for assessment and activity (AVT)*. Bergen: Centre for the Science of Learning & Technology (SLATE). Retrieved from https://bora.uib.no/handle/1956/20187

Acknowledgements

The authors wish to thank the technical experts that participated in the user testing for their valuable contributions.

Funding

This research has been supported by a University of Bergen PhD stipend at SLATE and funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 856954.