EdTech oder AdTech?

Untersuchungen zu Problemen datengetriebener Bildungsapplikationen

zerforschung

1 Einleitung

Schule wird digitaler – und damit kommt auch immer mehr Software zum Einsatz. Zwar wäre es wünschenswert, dass sich Apps und Programme im Bildungsbereich ausschließlich auf die Förderung des Lernens konzentrieren. Allerdings spielen bei derartigen Programmen auch immer wieder bestimmte Geschäftsmodelle eine Rolle, in deren Mittelpunkt Daten(sammlung) und Datenauswertung (z.B. für Werbung) stehen. Gerade im Bildungsbereich, in dem es um die hochgradig sensiblen Daten von Schüler*innen geht, erscheinen derartige Geschäftsmodelle extrem problematisch. Gleichzeitig ist es für Nutzer*innen in der Regel kaum nachvollziehbar, welche Daten gesammelt und wofür sie genutzt werden. So bleibt ihnen meist nichts anderes übrig, als sich auf die (Datenschutz-)Versprechen der Hersteller zu verlassen.

In diesem Beitrag wollen wir beispielhaft zeigen, wie wir rekonstruiert haben, wie sich Lern- und Schulapps technisch verhalten und welche Daten sie sammeln und übertragen. Denn auch wenn der Programmcode von Apps und anderer Software oftmals nicht einsehbar ist, können bestimmte technische Analysen darüber Aufschluss geben, wie die Software funktioniert. Dieses Vorgehen fasst man unter dem Begriff *Reverse Engineering* zusammen. In diesem Beitrag wollen wir verschiedene Herangehensweisen vorstellen, wie man grundsätzlich mittels *Reverse Engineering* Datenströme von Apps nachvollziehen kann.

Dazu möchten wir aber vorab noch eine deutliche Warnung loswerden: Wir können und wollen hier nur einen sehr kleinen Ausschnitt der Möglichkeiten abbilden. Dieser eignet sich für einen allerersten Einblick in Arbeits-

weisen der IT-Sicherheitsforschung – kann aber natürlich nicht einige Jahre Erfahrung in diesem Bereich sowie einen fundierten Wissenserwerb ersetzen.

Vorab möchten wir jedoch in wenigen Worten beschreiben, wer wir sind.

2 Das Kollektiv zerforschung

Wir sind das ehrenamtliche Kollektiv zerforschung. Wir beschreiben uns gerne selbst als ›ein freundliches Kollektiv aus Menschen, die Spaß daran haben, Technik auseinanderzunehmen, um zu verstehen, wie diese funktioniert‹.[1] Wobei wir streng genommen leider besonders häufig beobachten, wie technische Systeme *nicht* funktionieren. Um daraus zu lernen, bloggen wir unregelmäßig über unsere Funde auf zerforschung.org.

Aus dieser Neugier haben wir im Rahmen unserer ehrenamtlichen Arbeit seit Anfang 2021 regelmäßig Sicherheitslücken in verschiedenen Apps und anderen digitalen Diensten gefunden – darunter Kommunikationsplattformen, Lieferdienste und Software für Arztpraxen und Corona-Testzentren. Diese Sicherheitslücken haben uns theoretisch immer wieder Zugriff auf teils privateste Daten von tausenden bis Millionen von Menschen erlaubt.

Um zu verhindern, dass diese Sicherheitslücken ausgenutzt werden können, melden wir sie umgehend an die Hersteller in einem Prozess namens *Responsible Disclosure* oder *Coordinated Vulnerability Disclosure*. Das heißt: Wir informieren den jeweiligen Hersteller mit einer detaillierten Beschreibung des Problems sowie den Auswirkungen und geben ihm die Möglichkeit, die Fehler zu beheben.

Sobald der Hersteller die Lücken geschlossen hat und seine Software (zumindest an dieser Stelle) sicher ist, können wir darüber öffentlich berichten.[2] In einem Blogpost beschreiben wir dann, wie wir die Lücke gefunden haben und welche Auswirkungen sie hatte. Hierdurch erhoffen wir uns, dass

– andere Softwareentwickler*innen aus den Fehlern lernen und solche vermeiden können,

1 https://zerforschung.org/forscherinnen/

2 Unter gewissen Bedingungen kann dies auch vorher schon geschehen, etwa wenn eine zur Behebung mehr als ausreichende Frist verstrichen ist oder es Anzeichen gibt, dass die Lücke bereits öffentlich ausgenutzt wird. Das versuchen wir aber zu vermeiden, wann immer es geht.

- Nutzer*innen der betroffenen Dienste so ermöglicht wird, von den Lücken zu erfahren. Denn die meisten Hersteller informieren die Betroffenen nicht oder nur sehr beschönigend;
- wir eine öffentliche Debatte darüber anstoßen, wie es zu diesen Sicherheitslücken kommen konnte und was diese und die betroffenen Dienste für die einzelnen Betroffenen und unsere ganze Gesellschaft bedeuten.

Damit wollen wir einen Beitrag zu einer informierten Öffentlichkeit leisten, um die Wechselwirkungen zwischen Technik und Gesellschaft demokratisch diskutieren zu können.

3 Schutz von Kindern und Jugendlichen als gesamtgesellschaftliche Aufgabe

Grundsätzlich ist es eine gesamtgesellschaftliche Aufgabe, Kinder und Jugendliche besonders zu schützen – auch im digitalen Raum. Dafür gibt es bereits verschiedene Institutionen und Maßnahmen.

Die Datenschutz-Grundverordnung (DSGVO)

Mit der DSGVO gibt es eine größtenteils sehr gute Rechtsgrundlage für Datenschutz in Europa und Deutschland – mit besonderem Augenmerk auf den Schutzbedürfnissen von Kindern und Jugendlichen. Damit bildet die DSGVO die hauptsächliche rechtliche Grundlage für die Datenverarbeitung in Deutschland und gilt selbstverständlich auch für Bildungsapps. Daher bietet sich ein Blick auf einige dort normierte Anforderungen an.

Im eigentlichen Gesetzestext spielt vor allem der Artikel 8 eine Rolle: Dieser regelt, wann Minderjährige der Datenverarbeitung zustimmen können. Zudem regelt Artikel 12, dass die Informationen zur Datenverarbeitung einer App (welche Daten werden erhoben, wie werden sie verarbeitet, welche Rechte haben Nutzer*innen?) insbesondere für Kinder leicht zugänglich, transparent und verständlich sein müssen.

Der für die hier vorgestellten Anwendungen wichtigste Teil der DSGVO findet sich jedoch nicht im Gesetzestext, sondern in den ihm vorangestellten Erwägungsgründen. Hiervon befasst sich Erwägungsgrund 38 mit dem nötigen Schutz der Daten von Kindern (Hervorhebung durch die Autor*innen):

> Kinder verdienen bei ihren personenbezogenen Daten **besonderen Schutz**, da Kinder sich der betreffenden Risiken, Folgen und Garantien und ihrer Rechte bei der Verarbeitung personenbezogener Daten möglicherweise weniger bewusst sind. Ein solcher besonderer Schutz sollte insbesondere die Verwendung personenbezogener Daten von Kindern für Werbezwecke oder für die Erstellung von Persönlichkeits- oder Nutzerprofilen und die Erhebung von personenbezogenen Daten von Kindern bei der Nutzung von Diensten, die Kindern direkt angeboten werden, betreffen. […]

Der Erwägungsgrund trifft den Kern dieses Artikels: Wenn selbst Erwachsenen teils nicht bewusst ist, welche Auswirkungen die Erfassung ihrer personenbezogenen Daten haben kann, trifft dies auf Kinder und Jugendliche nochmal in besonderem Maße zu. Daher müssen Dienste, die sich an Kinder und Jugendliche richten, bei der Datenverarbeitung besonders sorgsam sein.

Erschwerend kommt hinzu, dass sich die Folgen einer derart umfassenden Sammlung der Daten von Kindern und Jugendlichen erst in einigen Jahren zeigen dürften. Denn so stark wie bei der aktuell schulpflichtigen Generation war das digitale Tracking durch Werbeanbieter und -netzwerke augenscheinlich noch nie.

Die Schulzeit ist zudem eine wichtige Entwicklungsphase, in der Schüler*innen ihre Persönlichkeit entwickeln – sich z. B. in ihrer politischen Überzeugung festigen oder die sexuelle und geschlechtliche Identität entdecken. Dies erfordert besonders geschützte Räume, auch um sich auszuprobieren und Fehler zu machen.

Wir wollen die Schule selbstverständlich nicht zu einem perfekten *geschützten Raum* verklären, in dem Schüler*innen sich gefahrlos entfalten können: Auch in Schulen finden sich zahlreiche Probleme wie Diskriminierung und Mobbing. Diese sollten gerade durch vermeintliche Bildungsapps nicht noch weiter verstärkt werden – etwa durch schlechte inhaltliche Moderation der Plattformen oder durch das Sammeln von Daten, die eine hohe Gefahr von Bloßstellung oder *Chilling Effects* mit sich bringen.

Anbieter von Bildungsapps haben auch deshalb eine hohe Verantwortung, weil den Schüler*innen oft keine Wahl oder Alternative zu ihrer Nutzung bleibt. Während andere datensammelnde Anwendungen erst das Vertrauen der Schüler*innen gewinnen müssen, haben Schüler*innen de facto keine Wahl mehr, sobald eine App im Unterricht zum Standard geworden ist: Sie müssen diese verwenden und die App kann die vergleichsweise schutzlo-

sen Schüler*innen als wertvolle, häufig wiederkehrende User*innen an ihre Werbekunden vermarkten.

Unsere Analyse unterschiedlicher Bildungsapps hat sehr deutlich gezeigt, dass Anbieter ihrer großen Verantwortung mitunter kaum gerecht werden. Stattdessen fanden wir unter anderem direkte Werbung in den Produkten, was dem eigentlich seit Jahren geltenden Grundsatz[3] ›keine Produktwerbung an Schulen‹ massiv widerspricht.

Die Datenschutzbehörden

In der Praxis reicht die reine Existenz der DSGVO nicht aus – sie muss auch durchgesetzt werden. Zuständig sind dafür die Datenschutzbehörden, die auf Bundes- und Länderebene angesiedelt sind und sich um die Einhaltung des gesetzlich verankerten Datenschutzes kümmern sollen. Im Rahmen unserer ehrenamtlichen Arbeit haben wir uns immer wieder mit Hinweisen oder der Bitte um Hilfe an verschiedene Datenschutzbehörden gewandt. Unsere Zusammenarbeit mit ihnen war meist sehr gut. Allerdings zeigt sich auch, dass die Behörden häufig erst anfangen zu arbeiten, wenn Probleme gefunden und an sie herangetragen werden.

Um das zu ändern, brauchen die zuständigen Behörden mehr personelle Kapazitäten, finanzielle Ressourcen und Kompetenzen. Nur so können sie ihrer Aufgabe nachkommen, verbreitete Software – und damit auch Bildungsapps – angemessen zu kontrollieren und bei Verstößen gegen die rechtlichen Vorgaben angemessene Strafen zu verhängen.

Die Schulbehörden und Schulen

Auch Schulbehörden und die Schulen selbst tragen substanzielle Verantwortung für die Sicherstellung des (Daten-)Schutzes von Kindern und Jugendlichen. Denn Schule ist eine Basisinfrastruktur unserer Gesellschaft – genauso wie auch digitale Schule, die nicht von privatwirtschaftlich getriebenen Akteuren übernommen werden darf.

Für eine digitale Basisinfrastruktur, die die Schüler*innen unterstützt und ihnen nicht schadet, braucht es jedoch einen entsprechenden Kompetenzaufbau und zusätzliche Ressourcen in den Schulen und den Schulbehörden. Ein*e

3 https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2013/Verbraucherbildung.pdf

Sportlehrer*in mit einer halben Stelle kann sich nicht nebenbei um die gesamte Schul-IT kümmern. Es braucht umfangreiches Wissen und Fähigkeiten, um Bildungsapps, ihre Funktionsweisen und ihre Geschäftsmodelle zu verstehen und kritisch zu hinterfragen.

Die Schüler*innen

Schüler*innen sind von ihrer Schule extrem abhängig, und damit auch von den digitalen Produkten, die in der Schule eingesetzt werden. Gleichzeitig wecken Schüler*innen große Begehrlichkeiten bei EdTech-Unternehmen. Dabei ist eine der obersten Aufgaben von Schulbehörden und Schulen, dafür zu sorgen, dass Schüler*innen im Schulkontext nicht zum Opfer datengetriebener Unternehmen werden. Genauso müssen die Schüler*innen aber auch selbst möglichst viel Wissen für einen kritisch-selbstbewussten Umgang mit Apps und Plattformen erlernen, um sich im weiteren Bildungskontext selbstständig für oder gegen den Einsatz von digitalen Produkten zu entscheiden.

Dennoch sollte hier die Entscheidungsfreiheit von Schüler*innen – auch im außerschulischen Bereich – nicht überschätzt werden: Wenn die Alternative zur Nutzung einer bestimmten, werbegetriebenen App lautet, durch schlechtere Vorbereitung eine schlechtere Note zu bekommen oder keinen Kontakt zu den Freund*innen zu haben, ist diese Entscheidungsfreiheit eher Illusion als Tatsache.

Die IT-Entwicklungsfirmen

Mit der steigenden Nachfrage nach Software im Bildungsbereich wächst auch das Angebot. Immer mehr Start-ups mit ganz verschiedenen Hintergründen wollen sich in dem Markt positionieren. Allerdings haben wir im Zuge unserer Recherchen verschiedene strukturelle Probleme bei den von uns analysierten Unternehmen festgestellt. In diesem Abschnitt stellen wir drei Probleme vor, die Kontrollen von Apps und Programmen im Bereich der Bildungsdigitalisierung in besonders großem Umfang nötig machen.

Zunächst zeigt sich bei den untersuchten Start-ups, dass vorgegebene Vorschriften wie die DSGVO oft sehr weitgefasst/großzügig ausgelegt bzw. scheinbar weniger ernst genommen werden. Entsprechende Start-ups rechtfertigen die nachlässig mitgedachte IT-Sicherheit mit angeblich fehlenden

Ressourcen.[4] Es ist jedoch wichtig zu betonen, dass sichere Software und Datenschutz keine optionalen Ergänzungen sind, die gegebenenfalls bei verfügbaren Ressourcen hinzugefügt werden können. Die DSGVO gilt auch für Start-ups. Wenn ein Produkt reif genug ist, um Nutzer*innendaten speichern zu können, dann muss es auch reif genug sein, diese für sich zu behalten.

Bei den untersuchten Start-ups zeigt sich eine bedenkliche Tendenz dazu, mögliche Konsequenzen für andere (bewusst) zu übersehen. Hier beziehen sich einige Gründer*innen von Start-ups immer wieder auf Silicon-Valley-Größen wie Facebook-Gründer Mark Zuckerberg als Vorbilder *(›The biggest risk you can take is not taking any risks‹)*. Inwieweit Unternehmen und/oder Gründer*innen, die wissentlich Nutzer*innen manipulieren, Hass und Hetze schüren und vieles mehr[5], eine Vorbildfunktion/Referenz für Start-ups sein sollten, bleibt jedoch fraglich.

Ferner gibt es bei Start-ups, die von sehr jungen Menschen gegründet werden, eine zusätzliche Herausforderung. In der Gesellschaft wird nach wie vor ein nahezu makelloser Lebenslauf angestrebt, der in der Regel Abitur, Auslandserfahrung und idealerweise Erfahrung in einem selbstgegründeten Start-up beinhaltet. Da viele Gründer*innen noch eng mit dem schulischen Umfeld verbunden sind, sehen sie hier oftmals eine niedrigschwellige Möglichkeit für einen beruflichen Einstieg in die Start-up-Welt. Es ist offensichtlich, dass die Gründung eines Unternehmens mit einer guten Idee und entsprechendem Fachwissen sehr bereichernd sein kann. Dabei kann man eine Menge lernen und sogar einen positiven Beitrag für die Gesellschaft leisten. Gleichzeitig muss jedoch klar sein, dass ein Unternehmen, das hauptsächlich aus karriereorientierten Gründen gegründet wird und dazu dient, den Lebenslauf zu verschönern, oft nicht die erforderlichen Standards erfüllen kann, um die sensiblen Daten von Kindern und Jugendlichen angemessen zu schützen. Dadurch verwandelt sich eine anfänglich spielerische Umgebung, in der auch Fehler gemacht werden können, um daraus zu lernen, in ein ernsthaftes Risiko für die betroffenen Nutzer*innen.

4 https://zerforschung.org/posts/scoolio/#aber-wir-sind-doch-ein-startup

5 https://netzpolitik.org/2021/vertuschte-skandale-facebook-ist-einfach-nicht-zu-trauen/

4 Einblicke in unsere Analysen von Apps im Bildungsbereich

In den vergangenen Jahren haben wir uns gezielt verschiedene Apps und Programme im Bildungsbereich angeschaut, weil diese große Versprechen machen: vom leichteren Lernen bis hin zum effizienteren Schulalltag für Schüler*innen und Lehrer*innen. Allen gemein ist, dass die Entwicklung von Apps wie auch der Betrieb von Servern und die permanente Pflege und Betreuung im Hintergrund aufwendig sind und finanziert werden müssen.

Wie genau das funktionieren kann, bleibt häufig unklar – denn in der Alltagserfahrung vieler Menschen sind Apps häufig kostenlos: im privaten Umfeld, weil die anbietenden Unternehmen sie als Serviceleistung für Kund*innen anbieten (z.B. Shopping- oder Mobilitätsanbieter), oder weil es eben um Daten als ›Preis‹ für die Nutzung geht, was in der Nutzung selbst jedoch erst mal nicht auffällt.

Im schulischen Rahmen ist insgesamt regelmäßig ungeklärt, wie Software finanziert werden soll. Bei Finanzierungspaketen wie dem ›DigitalPakt Schule‹ stand bislang die Anschaffung von Hardware im Vordergrund. Im Gegensatz dazu gibt es für Apps und Programme kaum flächendeckende Regelungen. Stattdessen sehen Software-Unternehmen einen Markt in der individuellen Nutzung durch Schüler*innen und Lehrer*innen. Daher versuchen solche Unternehmen, Lehrer*innen und Schüler*innen über informelle Wege (z.B. Direktakquise über Social Media) zu erreichen. Wir erachten es als grundsätzlich problematisch, dass es durch das Machtgefälle zwischen Lehrkräften und Schüler*innen regelmäßig dazu kommt, dass bestimmte Apps oder Plattformen von Lehrkräften im Unterricht eingeführt werden, ohne dass die Schüler*innen eine Wahl haben. Aber genauso erleben Lehrkräfte selbst durch die Intransparenz der Systeme ein strukturelles Machtgefälle gegenüber den Unternehmen.

Konkret haben wir uns im Rahmen unserer Analysen im Herbst 2021 drei Apps angeschaut: Scoolio, Learnu und StudySmarter. In allen dreien hatten wir zu dieser Zeit schwere Sicherheitslücken gefunden, die uns weitreichende Zugriffe auf bei den Appanbietern gespeicherten Daten erlaubten. Im Folgenden wollen wir diese Analysen genauer erläutern.

Zuvor stellen wir jedoch vor, wie wir die Methode des *Reverse Engineering* häufig anwenden.

5 Datenströme nachvollziehen: Ansätze des *Reverse Engineering*

Auf den ersten Blick sieht eine App für uns genauso aus wie für alle anderen auch: lustige Katzenbilder, Videos, Grafiken, Texte und bunte Buttons. Um zu verstehen, was sich dahinter versteckt, gibt es zwei Möglichkeiten:

- Wenn der Quellcode öffentlich zugänglich ist, können wir ihn lesen und so die technische Funktionsweise nachvollziehen.
- In den allermeisten Fällen ist der Quellcode aber nicht zugänglich, sondern der Zugang ist auf den für Menschen nicht lesbaren Binärcode reduziert. In diesem Fall bleibt uns nur, verschiedene Ein- und Ausgaben des Programms anzuschauen, zu experimentieren und über diese Methode sukzessive zu versuchen, die Funktionsweise so gut wie möglich zu rekonstruieren.

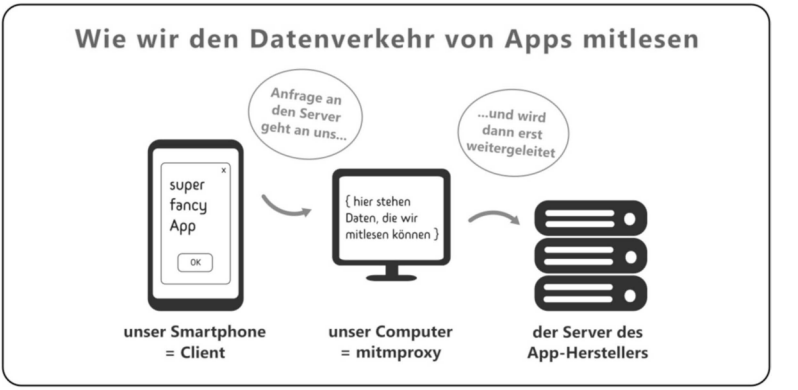
Die Vorgehensweise ist dabei meist ähnlich. Denn grundsätzlich funktionieren Apps so, dass sie auf einem Endgerät, z.B. einem Smartphone, einem Tablet oder einem Computer, installiert werden. Während der Nutzung sieht es zwar so aus, als passiere alles auf dem eigenen Gerät – immer wieder tauschen Apps aber Daten mit Servern aus.

Wenn wir nicht in den Code schauen können, um nachzulesen, wie App und Server miteinander kommunizieren, haben wir eine weitere Möglichkeit, das selbst herauszufinden: Wir schalten uns mittels eines sogenannten *Machine-in-the-Middle-Proxy* dazwischen. Dadurch können wir den Datenverkehr zwischen App und Server mitlesen.

Dass jemand den Datenverkehr zwischen App und Server mitliest, ist normalerweise unerwünscht. Deshalb überprüft eine App, ob der Server ein gültiges Zertifikat hat, sich also ausweisen kann. Ein solches Zertifikat können wir natürlich nicht vorweisen, auf unserem eigenen Handy können wir der App jedoch sagen, dass sie unser ungültiges Zertifikat dennoch akzeptieren soll.

Gleichzeitig suchen wir im Code der App nach weiteren hilfreichen Informationen. Hierfür experimentieren wir, wie oben beschrieben, mit dem Binärcode. Auf diese Weise lässt sich häufig – mit viel Zeit und Aufwand – rekonstruieren, was dieser Code genau tut.

Abb. 1: Wie das Mitlesen von Daten funktioniert (https://zerforschung.org/p/learnu/mitmproxy.png.)



Statt dieser umständlichen Rekonstruktion verwenden wir jedoch meist den simpelsten Ansatz des *Reverse Engineering*: Wir suchen nach menschenlesbaren Stellen, in denen beispielsweise Worte wie ›Admin‹ stehen oder an denen sich Links finden lassen, die man im Browser öffnen kann. Da wir wissen, wie Webentwicklung grundsätzlich funktioniert, können wir nun begründete Vermutungen anstellen, z. B. gibt es häufig Unterseiten mit Namen wie /login, /register, /docs oder auch /admin. Dort probieren wir verschiedene Eingaben aus und schauen uns an, wie der Server reagiert, ob er uns interessante Informationen oder auch Fehlermeldungen zurückgibt.

Immer wieder stoßen wir dabei auf Codes, Authentifizierungsschlüssel oder Ähnliches, mit denen wir uns dann dem Server gegenüber wiederum als ›legitime Nutzer*innen‹ ausgeben können und so auf große Datenmengen stoßen, die mit geringem technischem Fachwissen ohne Probleme abrufbar sind.

Sobald wir grundlegend verstanden haben, wie die App mit dem Server kommuniziert, brauchen wir die App nicht mehr, sondern können andere (Computer-)Programme nutzen, um mit dem Server zu ›sprechen‹. Technisch passiert im Hintergrund das Gleiche: Sowohl die App als auch wir mit unserer Software sprechen eine sogenannte Programmierschnittstelle (API) an, die uns dann Antworten schickt. Zum Beispiel: ›Bitte zeige mir mein persönliches Profil‹ – woraufhin der Server dann die entsprechenden Informationen wie beispielsweise Nutzername und E-Mail-Adresse liefert, die in der App oder auch in unserer Software angezeigt werden. Dabei sollte der Server nicht jede

beliebige Anfrage zulassen, sondern prüfen, welche Anfragen erlaubt sind und welche nicht. Beispielsweise sollte der Server so programmiert sein, dass er einer Nutzerin nur erlaubt, das eigene Profil zu verändern, nicht aber das von anderen Nutzer*innen.

Einblick I: Die App Learnu[6]

Die App Learnu – eine Zusammensetzung aus den Begriffen ›lernen‹ (›learn‹), ›verdienen‹ (›earn‹) und ›du‹ (›u‹) – bot einen Marktplatz für Hausaufgaben an: Schüler*innen sollten gelöste Hausaufgaben hochladen können, damit Credits verdienen und diese dann gegen Gutscheine, z.B. für große Online-Händler, tauschen können. Zusätzlich gab es ein Forum zum Austausch. Wer selbst keine Hausaufgaben hochladen wollte, konnte entweder ein Abo abschließen oder Werbung anschauen, um Zugang zu den fertig gemachten Hausaufgaben zu erhalten. Dieses System verstärkte damit bereits bestehende Ungleichheitsstrukturen: Wer Geld hatte, bekam relativ einfachen Zugriff auf hilfreiche Informationen. Wer kein Geld hatte, musste eher sein Datenvolumen aufbrauchen und Zeit investieren, um Werbung zu schauen – Zeit, die man auch mit Lernen verbringen könnte.

Wie sie in ihrem Podcast erzählen, waren die beiden Gründer von Learnu selbst noch in der Schule, als sie die Idee für ihre eigene App hatten. In ihrem Podcast räumen sie selbst ein, dass sie kaum Erfahrung mit App-Entwicklung hatten. Die App ließen sie extern entwickeln. Learnu kam bei der geplanten Zielgruppe gut an und wurde insgesamt von mehr als 500.000 Schüler*innen genutzt.

Um Sicherheitslücken bei Learnu zu finden, mussten wir die Learnu-App gar nicht erst starten. Stattdessen haben wir beim Anschauen des Binärcodes die URL admin.learnuapp.net/api gefunden.

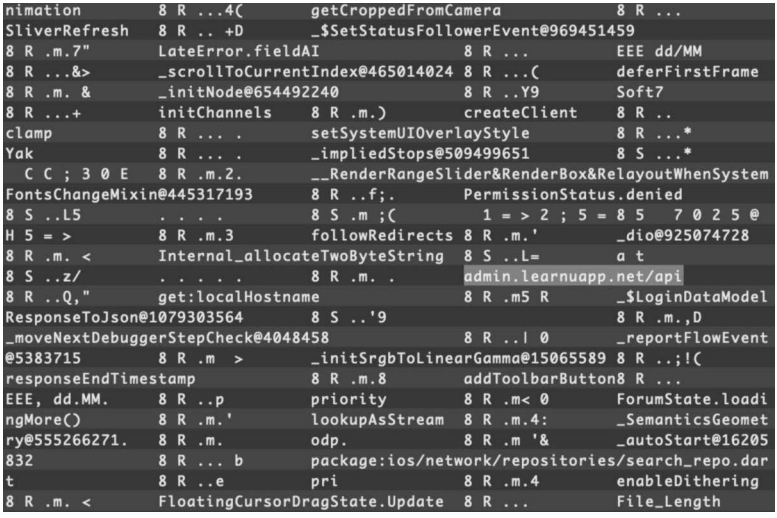
Nach etwas Herumprobieren fanden wir dann unter https://admin.learnuapp.net/docs eine Dokumentation, wie die Schnittstelle des Servers (API) funktioniert. Dort beschrieben war unter anderem ein Such-Endpunkt, der zu jeder dreistelligen Buchstabenkombination alle passenden Profile ausgibt. Suchen wir z.B. nach ann[7], bekommen wir die Daten der Nutzer*innenprofile mit ›Ann‹ im Namen, z.B. ›Anna‹ oder ›Anna-Marie‹. Dazu erhielten wir Informationen zum vollständigen Namen, die E-Mail-Adresse, das Datum der

6 https://zerforschung.org/posts/learnu/

7 https://admin.learnuapp.net/api/users/search/list?username=ann

Account-Erstellung, eine Selbstbeschreibung, IDs von Stadt und Schule sowie zur Aktivität auf der Plattform: Ist der Account noch aktiv? Ist ein Profilbild vorhanden? Wie viele Inhalte, Forenbeiträge oder Bestätigungen gab es?

Abb. 2: Binärcodes Learnu (https://zerforschung.org/p/learnu/bin%C3%A4rcode.png)



Der Server von Learnu begrenzte uns auf eine Suche pro Sekunde. Somit war es innerhalb von fünf Stunden theoretisch möglich, sämtliche auf der Plattform existierenden Profile abzurufen.[8] Durch eine weitere Sicherheitslücke waren zudem folgende Daten aller Nutzer*innen zugänglich:

- Vor- und Nachname, die bei der Registrierung als Klarname abgefragt werden
- Username innerhalb der App
- E-Mail-Adresse
- Stadt und Schule
- alle gestellten Fragen
- alle abgerufenen Antworten.

8 Es gibt $26^3 = 17.576$ Kombinationen von 3 Buchstaben (17.576 Sekunden sind etwas weniger als 5 Stunden).

Einblick II: StudySmarter

Das 2017 gegründete Unternehmen StudySmarter entwickelt eine App, die Schüler*innen und Studierende mit Lernkarten, Übungsaufgaben etc. bei der Prüfungsvorbereitung unterstützen möchte. Dafür wurden sie 2019 beim Branchen-Gipfeltreffen ›EdTechX‹ sogar zu ›Europas bestem EdTech-Startup‹ gekürt.[9]

Für die technische Analyse haben wir durch einen *Machine-in-the-Middle-Proxy* die Datenströme der App beobachtet, während wir uns registriert, eine Schule eingetragen und ein paar Lernkarten und Übungsaufgaben ausprobiert haben. Durch den Proxy konnten wir sehen, dass die App die Daten unseres Users über die URL https://prod.studysmarter.de/users/2983913 abrief. Wir zählten die in der URL angegebene Zahl um eins herunter und führten den Abruf erneut aus. Als Antwort schickte uns der Server das Profil einer anderen Person. Auch hier hätten wir in kurzer Zeit die Stammdaten aller rund drei Millionen registrierten Nutzer*innen abrufen können: E-Mail-Adresse, Schule/Universität, Studienrichtung, Geburtsdatum, Stadt, Bundesland, Land, Profilbild, ECTS-Punkte sowie ein Zugangstoken.

Mit diesem Zugangstoken war es wiederum möglich, den Account der Person zu übernehmen und so nicht nur Übungsaufgaben im Namen der Person zu erledigen, sondern auch die Daten der Person aus der App abzurufen, beispielsweise den Lernfortschritt – inklusive aller Details, wie eventueller Lernschwächen.

Einblick III: Scoolio

Auch das 2016 gegründete Scoolio wendet sich auf den ersten Blick mit diversen Werkzeugen für den Schulalltag dezidiert an Schüler*innen – vom Hausaufgaben-Planer über Notenübersicht und Klassenchat bis hin zur Nachhilfe-Vermittlung. Doch bei genauerer Betrachtung ist das Geschäftsmodell primär auf Scoolios Werbekunden ausgerichtet – auch weil es im Schulumfeld vergleichsweise schwierig ist, die Entwicklungskosten zu refinanzieren. Vermutlich wollen nicht allzu viele Schüler*innen ihr Taschengeld ausgerechnet für eine Schulapp ausgeben.

Die Anschubfinanzierung für die Entwicklung von Scoolio war jedoch eine ganz andere: Die Firma sammelte von 2016 bis 2021 mehr als zwei Millio-

9 https://www.studysmarter.de/presse/edtechx/

nen Euro von Investoren ein.[10] Durch eine Kleine Anfrage der Linken im Sächsischen Landtag[11] kam heraus, dass davon fast 1,4 Millionen Euro direkt aus staatlichen Mitteln stammen. Dazu kamen noch Mittel sächsischer Sparkassen.

Damit Scoolio gegenüber den Werbetreibenden das Versprechen zielgruppengerechter Werbung einlösen kann, muss die App möglichst viele Daten über Schüler*innen erfassen. Außerdem ist es für Scoolio attraktiv, die jungen Nutzer*innen so lange wie möglich auf der Scoolio-Plattform verweilen zu lassen, um ihnen möglichst viel Werbung auszuspielen.

Dafür haben die Entwickler*innen zentrale Elemente großer Plattformen nachgebaut – z.B. ein ›Tinder für Kinder‹: Dabei können sich Schüler*innen in derselben Schule oder Klasse finden, liken und über einen Chat anschreiben – von der ersten Klasse bis zum Schulabschluss. Eine Zugangsbeschränkung gibt es nicht.

Neben direkten Chats gibt es Klassen- und Schulchats sowie themenorientierte Chatgruppen bzw. ›Räume‹, die Schüler*innen selbst anlegen können. Viele davon sind so spezifisch benannt, dass schon die Mitgliedschaft in diesem Raum ein besonders schützenwertes Datum nach Art. 9 DSGVO ist. Hier eine kleine Sammlung von Chaträumen nach Kategorien besonders schützenswerter Daten (Zuordnung nach DSGVO):

- politische Meinung: ›Wir gegen Rassismus‹
- religiöse oder weltanschauliche Überzeugung: ›Muslime‹
- Gesundheitsdaten: ›abnehmen‹
- Sexualleben oder sexuelle Orientierung: ›LGBTQ+‹

In den Gruppen schien Scoolio nur sehr eingeschränkt zu moderieren, was besonders problematisch erscheint, weil die Chaträume derart einfach zugänglich sind. Unter den größten öffentlichen Räumen finden sich viele Räume wie ›Verliebt euch‹, ›Grube für Singles‹, ›Suche Freund zwischen 12 und 13‹ und ›nur Mädchen bis 10‹. Dabei hat Scoolio während unserer Analyse keinerlei Prüfung der eingegebenen Daten der User*innen durchgeführt oder den Zugang zu den Gruppen beschränkt. In einem Test haben wir einen Account mit dem

10 https://pitchbook.com/profiles/company/223999-03#signals

11 https://edas.landtag.sachsen.de/viewer.aspx?dok_nr=8094&dok_art=Drs&leg_per=7&pos_dok=&dok_id=277248

angeblichen Alter von 33 Jahren angelegt. Auch damit konnten wir allen diesen Gruppen beitreten, ohne von Scoolio-Moderator*innen entfernt zu werden.

Außerdem versucht die Plattform – dem klassischen Ansatz der Plattformökonomie folgend – die jungen Nutzer*innen mit immer mehr Features zu längerer Nutzung zu animieren und durch diese Features auch mehr Informationen abzufragen, etwa durch Persönlichkeitstests in Form von Job-Quizzes oder anderen lustigen Mini-Games.

Die Ergebnisse der Persönlichkeitstests wurden außerdem auch direkt an Arbeitgeber*innen zur sogenannten Leadgenerierung verkauft.[12] So steht dann im verkauften Datensatz neben Namen, Klasse und Kontaktdaten der Persönlichkeitstyp, z. B. ›Zielstrebiger Senkrechtstarter‹.[13]

Neben der massiven Datensammlung konnten wir in unserer Analyse zeigen, dass es auch im Fall Scoolio mit Grundkenntnissen in Web-Entwicklung oder IT-Sicherheit möglich war, auf hunderttausende Nutzer*innendatensätze zuzugreifen.[14]

6 Was nach unseren Analysen passierte

Bei unserer Meldung der Sicherheitslücken an die Unternehmen zeigten sich enorme Hürden. Beim Unternehmen Learnu war zunächst gar kein Kontakt für etwaige Sicherheitsprobleme auffindbar. Auf unsere Meldung über den generellen Kontaktweg kam zunächst keine Reaktion. Erst nachdem wir wiederholt nachfragten und teilweise im 15-Minuten-Takt anriefen, reagierte das Unternehmen und schaltete den Dienst dauerhaft ab. Damit waren die Datenschutzprobleme konsequent gelöst, jedoch stellte es die Nutzer*innen vor ein neues Problem: Wer den Dienst tatsächlich nutzte, um darüber zu lernen, stand unvermittelt allein da.

Auch im Fall StudySmarter haben wir keine Kontaktmöglichkeit für Sicherheitsanliegen gefunden. Also kontaktierten wir das Unternehmen über die allgemeine Kundensupport-Adresse sowie die Adresse für Datenschutzanliegen, auch hier zunächst ohne Erfolg. Als nächsten Schritt riefen wir die

12 Als Leadgenerierung (dt. Interessentengewinnung) bezeichnet man im Marketing das Finden von vielversprechenden Personen, die dann gezielt angesprochen werden können.

13 https://www.youtube.com/watch?v=PuSE5DW2sYc

14 https://zerforschung.org/posts/scoolio/

einzige auf der Website auffindbare Telefonnummer an: den Head of Sales, der unseren Bericht aus dem Spam-Ordner (wo er gelandet war) fischte und intern weiterleitete. Nachdem unsere Hinweise endlich die richtige Stelle bei StudySmarter erreicht hatten, wurde die Lücke dann tatsächlich in weniger als einer Stunde geschlossen. Außerdem hat StudySmarter direkt analysiert, wer auf die Daten zugegriffen hat, und ein ›Sicherheitsupdate‹[15] im Firmenblog veröffentlicht.

Insgesamt zeigen beide Beispiele, dass die Meldung der Sicherheitslücken sehr umständlich war und wir einige Hürden überwinden mussten, um die zuständigen Personen zu erreichen. Das muss nicht so sein: Über eine security.txt-Datei können Unternehmen einen Sicherheitskontakt und weitere nützliche Informationen für IT-Sicherheitsforscher*innen bereitstellen. Diese einzurichten dauert nur wenige Minuten, spart Sicherheitsforscher*innen wie uns jedoch viel Zeit und stellt vor allem sicher, dass Probleme möglichst schnell ihren Weg zur richtigen Ansprechperson finden.

Zwar wurden die konkreten von uns gefundenen Sicherheitslücken geschlossen, allerdings verdeutlichen sie ein darüber hinausgehendes Problem: Für sichere Software müssen Entwicklungsprozesse insgesamt so gestaltet werden, dass keine unbemerkten Fehler auftreten, die gravierende IT-Sicherheitslücken zur Folge haben. Dabei sollte das Prinzip gelten: Wenn eine Software marktreif genug ist, um Nutzer*innendaten zu speichern, dann muss sie auch reif genug sein, diese für sich zu behalten. Das muss durch angemessene Prozesse dauerhaft sichergestellt werden.

Daher bieten insbesondere punktuelle Sicherheitsüberprüfungen – wie sie in Reaktion auf unsere Analysen teils gefordert wurden – für sich allein keinen ausreichenden Schutz. So berichtete etwa StudySmarter, dass nur einen Monat, bevor die Lücke entstand, ein Sicherheitsaudit der Software durchgeführt wurde. Dass es trotzdem innerhalb von Minuten möglich gewesen wäre, auf private Daten von Millionen Nutzer*innen zuzugreifen, zeigt: Sicherheitsaudits alleine machen noch keine sichere Software. Denn einerseits können auch bei solchen Audits Fehler übersehen werden. Zum anderen bilden sie stets nur einen momentanen Stand der Software ab und schützen nicht davor, dass mit dem nächsten Update neue Lücken eingebaut werden.

15 https://www.studysmarter.de/magazine/studysmarter-sicherheitsupdate-november-2021/

7 Fazit

Wie wir in diesem Beitrag gezeigt haben, kann die Methode des *Reverse Engineering* wertvolle Beiträge zur Beurteilung von Lern- und Schulapps leisten. Neben der rein technischen Betrachtung von solchen Anwendungen lohnt auch ein genauerer Blick in die Finanzierungsstruktur eines Anbieters. Ein Mischmodell, also das Angebot eines kostenlosen Teils (z. B. Cloudnutzung) in Kombination mit werbefinanzierten anderen Services, kann dabei besonders problematisch sein, weil die Grenzen zwischen beiden Modellen im Alltag schnell verwischen bzw. sich *Lock-in*-Effekte einstellen können: Wo einmalig Zeit und Mühe in die Einrichtung einer digitalen Infrastruktur investiert wurde, fällt es schwer, diese wieder abzuschaffen – obwohl sie möglicherweise den nutzenden Kindern und Jugendlichen schadet.

Die Verantwortung für die IT-Sicherheit von EdTech-Produkten liegt primär bei den Herstellern. IT-Sicherheit darf dabei kein Add-on sein und nicht nur rein technisch verstanden werden. Software muss von Anfang an sicher und datensparsam konzipiert sein.

Ein kleiner Baustein dabei – der für Bildungsapps Standard sein sollte – sind sogenannte *Penetration Tests*. Wie bereits oben beschrieben, wird dabei im Auftrag des Herstellers nach Sicherheitslücken in der App gesucht. So sollen diese gefunden und behoben werden, bevor die App veröffentlicht wird.

Insgesamt zeigten unsere Analysen erhebliche Mängel in Sachen Datensicherheit, die auf teils drastische Interessenkonflikte zwischen Unternehmenszielen und gesetzlichen Bestimmungen zum Kinder- und Jugendschutz hinweisen. Anders ausgedrückt: Auch wenn ein Unternehmen behauptet, im Sinne der Schüler*innen zu agieren oder datenschutzkonform zu sein, sollte dies stets kritisch überprüft werden. Sichere und gute staatliche Plattformen müssen als Grundversorgung gelten und nicht als individuelles ›Nice-to-have‹. Denn Schule ist ein fundamentaler Teil der öffentlichen Daseinsvorsorge und muss vor Einflüssen und Interessen privatwirtschaftlicher Unternehmen in besonderem Maße geschützt werden.

Zentrale Take-Aways:

- Grundsätzlich sind die Daten von Kindern und Jugendlichen besonders schützenswert, was in der DSGVO klar verankert ist.
- *Reverse Engineering* ist eine Methode, um Datenströme in EdTech sichtbarer zu machen – wenngleich sie einige IT-Grundkenntnisse erfordert, die aber erlernbar sind.
- Generell sollte ein besonderes Augenmerk darauf liegen, welche Daten durch ein EdTech-Produkt gesammelt werden: Welche strukturierten Informationen wie Name, Alter, Schule und Klassenstufe werden abgefragt? Und zu welchen inhaltlichen Aussagen über sich selbst werden Schüler*innen beispielsweise über Chat- oder Forenelemente animiert?
- Werbe- und datengetriebene Geschäftsmodelle und Kindeswohl stehen in Konflikt zueinander. Daher sollte bei Anschaffung und Nutzung eines EdTech-Produktes dessen Geschäftsmodell immer mitdiskutiert werden. Auch Mechanismen des Produktes, die Schüler*innen zur Preisgabe von (persönlichen) Daten auffordern, sowie solche, die Schüler*innen zu einer langen Verweildauer im Produkt animieren, sind wichtige Alarmsignale für Praktiker*innen.