Datensparsame Webkonferenzplattformen

Dipl.Phys. Meinrad Rombach, Forum Agile Verwaltung e.V.

Über welche Art von Diensten reden wir, wenn wir von „Webkonferenz-Plattformen“ sprechen?
Welche Daten werden wie und zu welchem Zweck verarbeitet?
Wo lagern diese Daten und wer kann darauf zugreifen?

Nähere Hinweise dazu gibt das **Verfahrensverzeichnis**, doch das kann bei der Vielfalt der Produkte auf dem Markt durchaus unterschiedlich ausfallen. Die meisten Webkonferenz-Plattformen sammeln heute umfassend personenbezogene Daten ein, und viele geben diese mehr oder weniger restriktiv weiter. Hier ist aus Sicht des Datenschutzes die Frage zu stellen, welchem Zweck diese Datenerhebungen im Einzelnen dienen und welche Rechtsgrundlage für die Speicherung besteht. Ginge es alleine um die Erforderlichkeit für die Durchführung einer Webkonferenz, dürfte Art und Umfang der erhobenen Daten nicht so stark variieren, wie dies zu beobachten ist. Wodurch entsteht also der „Datenreichtum“ vieler, meist kommerzieller Plattformen?

Das eigentliche Konferenzsystem, das die synchrone (zu einem Termin stattfindende) Kommunikation durchführt, wird bei den meisten Anbietern ergänzt durch einige asynchrone Zusätze zur Erweiterung der Funktionalität. Meistens geht es darum, die Einstiegsschwelle in eine positive Ersterfahrung zu senken und die User danach mit all-inclusive Komfort bei der Stange zu halten. **Für diese Art Geschäftsmodell bieten viele Plattformen nicht nur die eigentliche Konferenz-Funktionalität, sondern auch eine Teilnehmer-Datenbank, Mailingdienste, Event-Marketing - gerne auch mit Besucher-Tracking,, Community-Funktionen für die Zeit zwischen den Konferenzen, Anwesenheits- und Aufmerksamkeits-Tracking oder ein Homepage-Baukasten für den unerfahrenen Webinar-Veranstalter**. Eine besonders wichtige Zusatzfunktion für gewerbliche Anbieter in einem Markt, in dem kostenlosen Einstiegsangebote üblich sind, ist die **Einbindung eines Online-Shops**, damit die Maschinenleistung effektiv monetarisiert werden kann. So verlangen einige Plattformen auch für die kostenlos angebotenen Schnupperdienste bereits die Eingabe von Zahlungs- und Adress-Daten, um später auf Knopfdruck schnell ein Geschäft abschließen zu können.

Für den personell aufwändigen technischen Support einer derart reichhaltig ausgestatteten "all-inclusive" Lösung ist es besonders kostengünstig, wenn bereits während der Sitzungen möglichst viele Details über die Hard- und Software sämtlicher Teilnehmergeräte eingesammelt werden. **Idealerweise (aus Sicht des Anbieters) wird bereits mit der Installation der beim ersten Einsatz herunter zu ladenden Software eine Fernwartungsmöglichkeit mit eingebaut**. In einem Fall ist dokumentiert, dass Code zum Erstellen von Bildschirmfotos und Übertragung der Systemdatenbank (Registry) des Teilnehmergeräts an den Hersteller vorhanden war, der ohne Kenntnis des Eigentümers agieren konnte. Eine solche vorsorgliche Datenerhebung zu potentiellen Supportzwecken ist dem Umstand geschuldet, dass diese Hersteller den Anwendersupport selbst durchführen, um Kunden entgegen zu kommen, die keine eigene IT-Support Infrastruktur haben. Auch bei dieser problematischen Zusatzfunktion in zwingend zu installierender Konferenzsoftware geht es um eine komfortable Senkung der Einstiegshürden die für Einsteuger kostenlos erscheint.

Eine datensparsame Fokussierung auf den eigentlichen Konferenzzweck bieten vor allem die Systeme, die dafür entwickelt wurden, sich in eine bestehende Infrastruktur des Betreibers zu integrieren und deshalb ohne eigenes datenintensives Management auskommen. Die Verfügbarkeit von Integrationsschnittstellen zu einer externen Teilnehmer- und Termin-Verwaltung ist ein Standardmerkmal der sogenannten „**Virtual Classroom**“ Lösungen, welche Videokonferenzen im Kontext von Lernplattformen bereitstellen.

Beispiele für solche Lösungen sind die bekannten Open-Source Projekte „JitsiMeet“ und „BigBlueButton“, die eine (sog. API-)Schnittstelle zur Verfügung stellen und bei entsprechender Konfiguration sehr datensparsam betrieben werden können.

Virtuelle Schulungsräume gibt es natürlich auch von kommerziellen Anbietern wie zum Beispiel „**Collaborate Ultra**“ des Universitäts-Ausstatters Blackboard. Das im angelsächsischen Raum auch als „CU“ bekannte Konferenzsystem wird meistens in akademische Lernumgebungen betrieben, integriert sich jedoch auch in andere bereits bestehende Teilhabe-Infrastrukturen, wo all die sensiblen Nutzer-Daten liegen und dort auch bleiben können. Diese akademische Plattform wurde hier als Beispiel ausgewählt, weil sie wie die beiden o.g. Open-Source-Projekte auf dem Webstandard webRTC basiert, bei dem keine potentiell problematische Installation von Software erforderlich ist.

Die für den Zweck „Webkonferenz mindestens erforderliche Datenerhebung wird im Folgenden **am Beispiel einer separat aufgesetzten CU-Konferenz** beschrieben, bei der das System als eigenständige Web-Konferenz mit Fokus auf Datensparsamkeit betrieben wird. In diesem Szenario ist das Konferenzsystem also nicht in ein Teilnehmerverwaltungssystem der Nutzerorganisation eingebunden, sondern wird als „managed hosting“ von einem externen Dienstleister betrieben.

Welche personenbezogenen Daten fallen bei einer puren Webkonferenz zwangsläufig an?
Dies läßt sich auch an der Frage festmachen:
Wie kann ich als Administrator mit maximalem Zugangsrecht zur Plattform im Nachhinein feststellen, wer an einer Veranstaltung teilgenommen hat und was er/sie im Verlauf beigetragen hat?

Während einer CU Sitzung entsteht auf dem Server eine Serie von Namen, die in **der Liste aller Anwesenden** angezeigt wird. Erfolgt der Zugang zur Sitzung über einen nicht personalisierten Gastlink , haben die Teilnehmenden diesen „Anzeigenamen“ beim Eintritt selbst eingegeben. Ob dies Pseudonyme sind oder echte Namen, das ist bei dieser Zugangsform ungewiss. Wer über den Gastlink Zugang hatte, dem stand es ebenso frei, einen Namen zu wählen, wie auch von dieser Liste der Anwesenden einen Screenshot zu machen, solange die Veranstaltung lief.

Der Server weiß in diesem Szenario „Gast-Zugang“ über den einzelnen Gast wenig, denn hier gibt es weder einen authentifizierten Identitätsprovider („Anmeldung mit Facebook, Google, etc.“) , noch wird die E-Mailadresse erfasst, zB mit der hilfsweisen Begründung „Wir erinnern Ihre Teilnehmer automatisch, dass es bald los geht“. Falls einzelne Gäste mal ausnahmsweise einen technischen Anlass sehen, Informationen über das eigene System an den CU Support zu übermitteln, kann das in transparenter Weise mit der Funktion "Problem melden" geschehen, wobei man die Log-Informationen auch mit Copy&Paste übernehmen kann, ohne dass der Hersteller davon erfährt.

Selbstverständlich existiert auch auf dem CU-Server ein Reporting, das die Liste der Anzeigenamen enthält. Diese Angaben sind über den administrativen Account einsehbar, der die entsprechende Konferenz generiert und den dabei erzeugten Gastlink erhalten und verteilt hat. Darüber hinaus gibt es noch einen übergeordneten Manager-Account auf dem System des Herstellers, der die Accounts von Administratoren anlegt. Mit dem Manager-Account können auch Event-übergreifende Reports abgerufen werden. Diese enthalten neben dem Anzeigennamen auch die Zeiten von Zugang und Weggang.

Eine Authentifizierung von einzelnen Teilnehmern über solche Accounts auf dem Konferenzsystem ist also nicht zwingend erforderlich. **Die Nutzung der administrativen Zugänge lässt sich deshalb z.B. auch an einen externen Dienstleister auslagern**, mit dem ein entsprechender Support-Vertrag zu schließen wäre, zusätzlich zum Auftragvserarbeitungsvertrag mit dem Hersteller Blackboard.

Diese Supportinstanz kann im Auftrag Räume erstellen mit entsprechenden Zugangslinks. Sie kann auch (vergleichbar mit einem externen Pförtnerdienst) als Treuhänder für die Herausgabe von Nutzer-Reports an Berechtigte dienen. Falls nichts dagegen spricht (typischerweise ein Betriebsrat oder eine fehlende Rechtsgrundlage) würde eine Anwesenheits-Liste kryptografisch gesichert vom Support-Dienstleister an den Auftraggeber übermittelt werden mit dem Hinweis, dass eine Weitergabe dieser Liste an interessierte Dritte ebenfalls eine Rechtsgrundlage benötigt .

Aufgabe dieser Support-Instanz kann bei entsprechender Beauftragung auch eine Pseudonymisierung der privilegierten Nutzerschaft sein. Das sind diejenigen Personen, die mit erweiterten Rechten im Konferenzraum tätig werden sollen. Für diese Einzelpersonen ist ein individueller Zugang erforderlich, der bei Zutritt zur Konferenz automatisch erhöhte Rechte einräumt. Für die Erstellung solcher individueller Zugänge fordert die CU Administrationsumgebung wie viele andere auch die Angabe von

- Anzeigename
- eMail-Adresse
- Rolle (z.B. Moderationsrolle)

Die eMail-Adresse benötigt der Hersteller, um dem entsprechenden Nutzer den Zugangslink direkt zusenden zu können.

Beim angenommenen Betrieb als Managed Hosting kann die zwischengeschaltete Supportinstanz an dieser Stelle wieder treuhänderisch tätig werden, indem zur Erzeugung von Moderationszugängen nicht die echten Mailadressen der privilegierten Benutzer eingetragen werden, sondern eine Adresse, die die Support-Instanz empfängt. Dadurch erlangt zwar die Support-Instanz Kenntnis der individuellen Zugangslinks, spart aber die Übermittlung von echten User-Mailadressen an den Server ein. Lediglich der Anzeigename, der die erweiterten Rechte erhalten soll, muß der Webkonferenz Datenbank mitgeteilt werden. Dies kann wieder ein Pseudonym sein zB „Dozent“ oder „Moderator“

Auf der CU-Plattform kann eine **Aufzeichnung der Veranstaltung** erstellt werden, bei der die Anzeigenamen nicht gespeichert werden, da sie bereits während der Sitzung anonymisiert wurden. So ist auch bei einer Weitergabe der Aufzeichnung keine direkte Zuordnung von Teilnehmernamen möglich, es sei denn dass das gesprochene oder geschriebene Wort einen Personenbezug herstellt.

Bei einer solchen datensparsam extern gemanagten Webkonferenz führt der einzige Weg zu den Kontaktdaten der Eingeladenen dorthin, wo das Event-Management erfolgt. Für privat organisierte Veranstaltungen wäre der Veranstalter für die Datenhaltung selbst verantwortlich, doch wie sieht das bei einer öffentlichen Einladung aus?

Stellvertretend für viele andere wird hier das Xing-Eventmanagement-System betrachtet, um den Einfluss sozialer Netzwerke mit abzubilden. Auf der Datenbank von Xing liegen Infos über diejenigen, die sich für die Webkonferenz angemeldet hatten. Diese Information wird dort für Xing-Mitglieder freizügig präsentiert, da dies das Kerngeschäft von Xing ist, in das alle eingewilligt haben. Ob eine angemeldete Person jedoch wirklich an der CU Session teilgenommen hat, weiß die Xing-Datenbank nicht, weil das Event-Management-System nicht mit dem Konferenzsystem verknüpft ist. Wenn der Veranstalter die Option einräumt, auch ohne Xing-Anmeldung an der Konferenz teil zu nehmen, dann ist für diejenigen, die dies wünschen, weiterhin eine pseudonymisierte Teilnahme möglich.

Bei einer in Xing datentechnisch integrierten Webkonferenz-Lösung wäre das anders. Hier wäre die Mitgliedschaft bei Xing für eine Teilnahme an der Veranstaltung obligatorisch, und die Identität der Gäste wäre automatisch mit den Metadaten ihrer tatsächlichen Teilnahme verknüpft. Dies gilt in gleicher Weise auch für alle Webkonferenz-Plattform, welche den anfangs beschriebenen „all-inclusive“ Komfort anbieten. Diese können grundsätzlich kein „Privacy by Design“ bieten.

Einsatzfelder einer Webkonferenz-Plattform mit Veranstaltungsmanagement

Ich hoffe, ich konnte mit dieser Darstellung deutlich machen, dass "soziale" Netzwerk-Plattformen mit Gewinnerzielungsabsicht, reichhaltigem Wissen über individuelle User-Aktivitäten und einem eigenem Webkonferenzdienst sich besonders eignen würden für das Einsammeln von Kontaktdaten. Der Event-Typ „kostenloses Webinar zum Thema X“ ist auch entsprechend beliebt, um Kontakte zu knüpfen zu Personen, die am Thema X interessiert sind.

Der Fokus eines virtuellen Klassenraums wie BigBlueButton oder Blackboard CU liegt dagegen in der Unterstützung von verteilt sitzenden Arbeits-/Lern-/Interessen-Gruppen, die sich bereits unter einem gemeinsamen organisatorischen Dach gefunden haben und auf Augenhöhe miteinander interagieren wollen, ohne dies gegenüber Außenstehenden zu veröffentlichen.

Wie dieses organisatorische Dach implementiert ist, sollte die Organisation selbst gestalten können, nicht der Webkonferenz-Provider. Diese Gestaltung erfordert eine gewisse Flexibilität, die von dem fixen Dach eines Webkonferenz-Providers nicht geliefert werden kann.

Eine Hochschule, in der sich junge Menschen für ein Studium einschreiben und in Folge Vorlesung, Seminar, Kolloquium oder Vortrag belegen, geht mit personen-bezogenen Daten anders um als eine Organisation, die eigene Mitarbeiterinnen und Mitarbeiter qualifizieren will, oder die projektbezogen agil kollaborierende Teams bildet, in denen auch Externe mitwirken. Ein auf Datensparsamkeit getrimmtes Konferenzsystem fügt sich hier besser ein als eine „Application-as-a-Service-Plattform“, die im Kern zu extern gelagerten Duplikaten personen-bezogener Daten führt, weil für die Teilnahme an Web-Konferenzen zwingend auch ein externe Web- Konferenz –Datenbank mit Nutzerdaten beliefert werden muss, was die Organisation eigentlich gar nicht benötigt.

Selbstverständlich gibt es einige denkbare Anwendungsszenarien, in denen vollständig ausgelagerte Verwaltung die beste Lösung ist. Gerade in der derzeitigen Krise waren solche funktionell reichhaltigen Plattformen hilfreich, weil sie keinerlei sonstige Infrastruktur bei der Nutzerschaft voraussetzen und auch Laien bei allen Schritten zur Organisation und Verwertung von Webkonferenzen an die Hand nehmen. Gleichwohl stellen solche naturgemäß besonders datenhungrigen Lösungen erheblich höhere Anforderungen an den Datenschutz, was einer schnellen Umsetzung einer langfristig tragfähigen Lösung sehr im Wege steht.

Zusammenfassung

Für den originären Verarbeitungszweck "Webkonferenz" ist eine Speicherung von Nutzerdaten nicht zwingend erforderlich, abgesehen von den üblichen Server-Logs beim Zugriff auf Webseiten und den zum jeweiligen Konferenzzweck inhaltlich übermittelten Daten, die nicht Gegenstand dieser Betrachtung waren. Die Bereitstellung von Anzeigenamen kann pseudonym erfolgen, wenn die Zugangsmethode Gastzugang verwendet wird. Erforderlich für den Betrieb sind nur die Angaben für ein administratives Konto, das die Webkonferenzen aufsetzt und Zugänge erzeugt.

Webkonferenzen, die auf einer Plattform zu betreiben sind, welche aufgrund von integrierten Zusatzfunktionen nur mit Speicherung von schützenswerten Nutzerdaten funktioniert, müssen insofern als nicht datensparsam gestaltet und deshalb abgewertet werden. Jede Auslagerung der Teilnehmer- und Terminverwaltung an eine Webkonferenz-Plattform stellt faktisch eine Erweiterung der Auftragsverarbeitung dar, die eine aus Sicht des Datenschutzes sinnvollere interne Verwaltung der Nutzerdaten externalisiert und dadurch grundsätzlich trennbare sensible Daten zusammen führt.