THOR — Ein Datenschutzkonzept für die Industrie 4.0

Datenschutzsysteme für die Smart Factory zur Realisierung der DSGVO

Christoph Stach[1], Frank Steimle[1], Bernhard Mitschang[1]

Abstract: Der Aufschwung des *Internets der Dinge* (*IoT*) sorgt für eine voranschreitende Digitalisierung. Sensoren in Alltagsgegenständen erfassen unterschiedliche Aspekte des täglichen Lebens. Durch eine Vernetzung dieser Geräte, lassen sich die Daten miteinander kombinieren und daraus neues Wissen generieren. In der *Industrie 4.0* werden beispielsweise die am Produktionsprozess beteiligten cyber-physischen Systeme dazu genutzt, um mit den von ihnen erfassten Daten Produktionsprozesse zu optimieren. Da auch der Mensch ein relevanter Bestandteil des Produktionsprozesses ist, werden z. B. mittels *Smart Watches* auch über diesen viele Daten erfasst. Nicht erst mit der Einführung der neuen *Datenschutzgrundverordnung* (*DSGVO*) sind hierbei allerdings Datenschutzanforderungen zu beachten: Es müssen nicht nur die privaten Daten der Nutzer geschützt werden, sondern es muss auch sichergestellt werden, dass die Datenverarbeitung und -analyse dadurch so wenig wie möglich behindert werden. Wir stellen hierfür ein neuartiges Da**t**ensc**h**utzk**o**nzept für die Indust**r**ie 4.0 (**THOR**) vor, mit dem Kompromisse zwischen erforderlichem Datenschutz und gewünschter Datenqualität gefunden werden können, der der DSGVO genügt.

Keywords: Datenschutz; Internet der Dinge; Sensoren; Industrie 4.0; Datenströme; Smart Devices.

1 Einleitung

Heutzutage sind in Alltagsgegenständen, wie Armbanduhren, eine Vielzahl an Sensoren verbaut. Diese *Smart Devices* erfassen damit Daten über ihre Nutzer[2], wodurch sie Leistungen anbieten können, die weit über ihren eigentlichen Verwendungszweck hinausgehen. So macht ein verbauter Pulssensor aus einer normalen Armbanduhr ein Herzfrequenz-Messgerät. Darüber hinaus lassen sich die Smart Devices untereinander vernetzen und so die Daten von unterschiedlichen Geräten verknüpfen. Die Idee dieses Internets der Dinge (engl. *Internet of Things*, *IoT*) ist es, mannigfaltige Daten kontinuierlich anzusammeln, umfangreiche Analysen darauf anzuwenden und dadurch neues Wissen zu generieren [Ha15].

Aufgrund dieser Daten – und dem damit verbundenen Potential – steigt der ökonomische Wert des IoTs stetig. Gartner prognostiziert, dass es bis zum Jahr 2020 über 26 Milliarden Smart Devices geben wird und sich damit ein wirtschaftlicher Mehrwert von 1,9 Billionen US-Dollar erzielen lässt. Insbesondere das Gesundheitswesen und die Fertigung profitieren

[1] Universität Stuttgart, Universitätsstraße 38, D-70569 Stuttgart, Vorname.Nachname@ipvs.uni-stuttgart.de

[2] Der Begriff „Nutzer“ beschreibt im Rahmen dieses Artikels die Person, die Smart Devices verwendet, d. h., über die Daten erhoben werden. Mit dem Begriff „Nutzer“ seien beide Geschlechter gleichermaßen adressiert.

vom IoT [Mi13]. Während sich IoT-Lösungen für das Gesundheitswesen stark auf den Patienten fokussieren und darauf die Behandlung zu erleichtern, adressieren IoT-Lösungen für die Fertigung primär die Erfassung und Nutzung von Maschinendaten, die während des Produktionsprozesses anfallen. Da Produktionsdaten keinen Personenbezug haben, werden in diesem Umfeld technische Datenschutzlösungen weitestgehend vernachlässigt [Kh17].

Immer mehr wird allerdings erkannt, dass auch in der *Industrie 4.0* der Faktor Mensch hochgradig relevant ist [Ka17]. Daher werden Möglichkeiten gesucht, die Werker mit Sensoren zu versehen und so deren Arbeit numerisch zu erfassen. Ein an *Bring Your Own Device* (*BYOD*) angelehnter Ansatz besteht darin, sämtliche Werker mit *Smart Watches* auszustatten, die sie auch privat nutzen dürfen [UA16]. Während es hinsichtlich der Datensicherheit für Smart Devices mehrere Konzepte gibt, die eine Isolation entweder auf Anwendungsebene [OB16] oder auf Datenebene [Ki17] einführen, stellt der Datenschutz insbesondere auf Smart Watches ein überwiegend unzureichend gelöstes Problem dar [St18a].

Spätestens mit dem Inkrafttreten der *Datenschutzgrundverordnung* (*DSGVO*) [EU16], besteht daher ein Nachbesserungsbedarf, um Sanktionen abzuwenden. Insbesondere müssen die Datenschutzlösungen auf den speziellen Anwendungsfall Industrie 4.0 angepasst sein. So müssen personenbezogene Daten zwar ausreichend geschützt werden, die Datenverarbeitung soll aber sowenig wie möglich eingeschränkt werden [Se17]. Dies wird am folgenden Beispiel deutlich: Eine einfache Form des Datenschutzes besteht darin, die Konnektivität der Smart Watches zu deaktivieren – dadurch können keine davon erfassten Daten weitergegeben und verarbeitet werden. Von diesem Vorgehen sind allerdings auch alle unkritischen Daten betroffen, die ebenfalls von Smart Watches erfasst werden.

Da aktuelle Datenschutzlösungen diesbezüglich zu restriktiv vorgehen und somit die Datenanalyse unnötig stark einschränken, führen wir mit **THOR** ein neuartiges Da**t**ensc**h**utzk**o**nzept für die Indust**r**ie 4.0 ein. Zu diesem Zweck erbringen wir die folgenden Forschungsbeiträge: FB_1 Wir stellen einen Industrie 4.0 Anwendungsfall vor. FB_2 Wir leiten Datenschutzanforderungen aus dem Anwendungsfall ab. FB_3 Wir stellen mit THOR eine Datenschutzlösung für die Industrie 4.0 vor, die sowohl die privaten Daten der Nutzer schützt als auch die Datenverarbeitung und -analyse so wenig wie möglich behindert. FB_4 Wir evaluieren THOR und zeigen, dass THOR die abgeleiteten Anforderungen erfüllt.

Der Rest dieses Artikels ist wie folgt gegliedert: In Abschnitt 2 wird ein Anwendungsfall im Bereich der Industrie 4.0 beschrieben, auf dessen Basis wir in Abschnitt 3 Datenschutzanforderungen ableiten. Abschnitt 4 befasst sich mit verwandten Arbeiten, während THOR Gegenstand von Abschnitt 5 ist. Abschnitt 6 evaluiert unseren Ansatz anhand der abgeleiteten Anforderungen, bevor Abschnitt 7 diesen Artikel zusammenfasst.

2 Anwendungsfall

Im Folgenden wird ein Industrie 4.0 Anwendungsfall beschrieben, aus dem klar ersichtlich wird, welche Implikationen die DSGVO mit sich bringt. Dabei wird das Beispiel

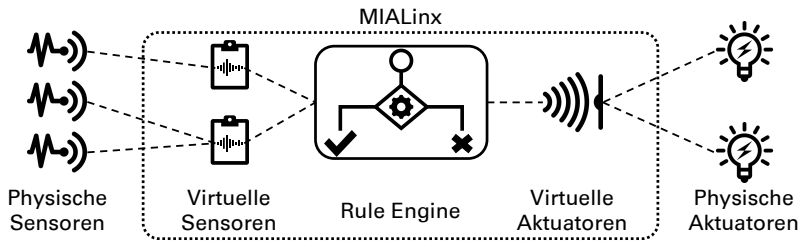

Abb. 1: Grundkonzept von *MIALinx*, einer Automatisierungsplattform für die *Smart Factory*.

der sogenannten *Smart Factory* herangezogen, also einer Produktionsumgebung, die sich weitestgehend selbst organisiert. Hierfür ist es nötig, dass die an der Produktion beteiligten Maschinen mittels Sensoren relevante Kennzahlen erfassen und selbstständig beispielsweise auf Abweichungen von Normwerten reagieren. Damit dies möglich ist, bedarf es eines Zusammenspiels von *Sensoren* (d. h., Komponenten, die Aspekte ihrer Umwelt erfassen können) und *Aktuatoren* (d. h., Komponenten, die ihre Umwelt beeinflussen können). *MIALinx* [Wi16; Wi17] ist eine Plattform, die eine solche Vernetzung in der Produktion ermöglicht. Abb. 1 zeigt das Grundkonzept von MIALinx. Jeder physische Sensor und Aktuator wird auf einen virtuellen Sensor (respektive Aktuator) abgebildet. Bei diesen virtuellen Komponenten handelt es sich um Software, die Rohdaten von den physischen Sensoren erfasst und in aufbereiteter Form anderen Software-Komponenten zur Verfügung stellt, bzw. Schnittstellen mit denen Aktionen auf physischen Aktuatoren ausgelöst werden können anbietet. Eine virtuelle Komponente kann dabei auch Werte von mehreren Sensoren erfassen bzw. mehrere Aktuatoren ansteuern. Die eigentliche Verknüpfung erfolgt über eine *Rule Engine*, die *WENN-DANN-Regeln* auswerten kann. Eine solche Regel könnte beispielsweise $S_1 > 42 \rightarrow A_2\left(N_3\right)$ lauten, d. h., wenn der Wert des Sensors S_1 den Schwellwert von 42 übersteigt, dann soll die Nachricht N_3 an den Aktuator A_2 geschickt werden. Dabei kommuniziert die Rule Engine ausschließlich mit den virtuellen Komponenten.

Während die Vision der Smart Factory gänzlich ohne menschliches Zutun auskommt, ist dies in der Realität nicht möglich. Für Wartungs- und Reparaturmaßnahmen oder in einem schwerwiegenden Fehlerfall ist beispielsweise die Expertise eines Werkers erforderlich. Die *soziale Fabrik* [Ka17] adressiert dies, indem sie ein soziales Netzwerk für die Industrie 4.0 einführt. In dem Netzwerk ist u. a. hinterlegt, welcher Werker welches Fachwissen besitzt. Dadurch ist es möglich im Fehlerfall automatisch alle Experten auf dem jeweiligen Gebiet zu identifizieren. Durch eine Integration dieses Netzwerks in MIALinx können die Regeln definiert werden, die bei einem Problem die Experten direkt informieren. In der Regel sollen allerdings nicht alle Experten kontaktiert werden, sondern nur derjenige, der das jeweilige Problem am schnellsten lösen kann, d. h., der Werker, der sich räumlich am nächsten befindet und aktuell keine anderen Termine hat. Damit MIALinx dies entscheiden kann, müssen auch die Werker mit Sensoren ausgestattet werden. Dies kann dadurch realisiert werden, dass der Arbeitgeber alle Werker mit Smart Watches ausstattet. Auf diesen Geräten können nicht nur Termine verwaltet werden, sondern die darin verbauten Beschleunigungssensoren

und Gyroskope gestatten ebenfalls eine sehr genaue Indoor-Ortung [Hs14]. Darüber hinaus lassen sich mittels Smart Watches auch die aktuellen Aktivitäten der Nutzer sehr präzise bestimmen [Sh17]. So kann beispielsweise erkannt werden, wenn ein Werker gerade an einer anderen Maschine eine Reparatur durchführt, ohne dies im Kalender eingetragen zu haben. In diesem Fall kann der Wartungstermin automatisch vermerkt werden. Es ist ebenfalls möglich, Werker über deren spezifische Bewegungsmuster eindeutig zu identifizieren [Co11]; somit kann ein eindeutiges Mapping von Smart Watch auf Person erfolgen.

An diesem Beispiel wird offensichtlich, dass in der sozialen Fabrik eine Vielzahl an personenbezogenen Daten erfasst wird. Ferner erfassen Smart Watches noch weitere, in höchstem Grad persönliche Daten, wie beispielsweise die Herzfrequenz und andere medizinische Daten [AA16]. Daher muss genau betrachtet werden, inwiefern sich dies mit der DSGVO vereinen lässt. So wird beispielsweise in Artikel 5, Absatz 1 (c) *Datenminimierung* gefordert. Angewandt auf das obige Beispiel würde dies u. a. bedeuten, dass sämtliche Gesundheitsdaten auf keinen Fall erfasst werden dürfen. Die *Zweckbindung* (Artikel 5, Absatz 1 (b)) bedeutet für die Aktivitätserkennung, dass lediglich erfasst wird, ob der Werker gerade beschäftigt ist, nicht aber, ob er z. B. in den Pausen raucht. Werden Daten aus Gründen des Datenschutzes allerdings verfälscht, so ist dennoch dafür Sorge zu tragen, dass die *Richtigkeit* der erhobenen Daten erhalten bleibt (Artikel 5, Absatz 1 (d)). Dies wirkt sich erheblich auf die Identifikation anhand von Bewegungsmuster aus, für den Fall, dass aufgrund der Zweckbindung bestimmte Muster nicht erfasst werden. Schließlich fordert die DSGVO auch *Transparenz* (Artikel 5, Absatz 1 (a)) und *Rechenschaftspflicht* (Artikel 5, Absatz 2), so dass ein Verifikationsmechanismus erforderlich ist, über den Betroffene sich informieren können, ob deren personenbezogene Daten ausreichend geschützt wurden.

3 Datenschutzanforderungen

Aus dem obigen Anwendungsfall lassen sich folgende Datenschutzanforderungen ableiten:

- **R_1 Datenschutz für Smart Devices.** Bestimmte Daten, die für den angestrebten Verarbeitungszweck nicht relevant sind müssen frühzeitig herausgefiltert werden, bevor sie in das Datenstromsystem (MIALinx) gelangen können. Dies muss daher direkt auf dem physischen Sensor erfolgen. Daher müssen technische Datenschutzmechanismen für Smart Devices bereitgestellt werden.
- **R_2 Einheitliche Konfiguration.** In der Industrie 4.0 kommen viele heterogene Smart Devices zum Einsatz. Die Konfiguration der erforderlichen Datenschutzmechanismen muss für die IT-Abteilung handhabbar sein. Daher muss es möglich sein, Konfigurationen nur einmal zentral zu definieren und anschließend automatisch auf die Geräte auszubringen.
- **R_3 Kontextabhängiger Datenschutz.** Da die Smart Devices von den Werkern auch in den Pausen getragen werden, muss der Datenschutzmechanismus kontextabhängig agieren. Beispielsweise müssen wesentlich restriktivere Regeln außerhalb der Arbeitszeit angewandt werden, die keinerlei Daten an den Arbeitgeber weitergeben.

- **R_4 Datenschutz für Datenstromsysteme.** Manche Daten müssen unverfälscht an das Datenstromsystem weitergegeben werden, damit die angestrebten Analysen durchgeführt werden können. Daher müssen auch an dieser Stelle Datenschutzmechanismen existieren, da erst hier alle Datenquellen und -verarbeitungsarten bekannt sind.
- **R_5 Musterbasierter Datenschutz.** Die Datenschutzmechanismen müssen z. B. in der Lage sein, das Aktivitätsmuster „rauchen“ zu verbergen, das Muster „reparieren“ aber offenzulegen. Diese Muster werden aber von den selben Sensoren erfasst und unterscheiden sich lediglich in der Sequenz der eingehenden Sensordaten.
- **R_6 Aufrechterhaltung der Datenqualität.** Die Datenschutzmechanismen müssen unterschiedliche Datenschutztechniken unterstützen. So kann jeweils die Technik ausgewählt werden, die den geringsten negativen Einfluss auf die Datenqualität hat, damit alle Verarbeitungen, die keine privaten Daten offenlegen, weiterhin möglich sind.
- **R_7 Verifikation.** Die Datenschutzmechanismen müssen den Nutzern Möglichkeiten anbieten, mit denen sie überprüfen können, ob die Mechanismen korrekt konfiguriert sind. Es muss daher nachweisbar sein, dass weder private Daten zur Verarbeitung freigegeben, noch unkritische Daten unnötigerweise zurückgehalten wurden.

4 Verwandte Arbeiten

Wie die Datenschutzanforderungen zeigen, müssen von einem Datenschutzkonzept für die Industrie 4.0 zwei heterogene Arbeitsweisen unterstützt werden, nämlich die Datenverarbeitung sowohl auf Smart Devices als auch durch Datenstromsysteme. Unseres Wissens nach gibt es aktuell keinen einheitlichen Mechanismus. Im Folgenden werden daher verwandten Arbeiten im Bereich der Smart Devices und der Datenstromsysteme separat betrachtet.

Datenschutzmechanismen für Smart Devices. Viele wissenschaftliche Arbeiten beschäftigen sich mit Datenschutzmechanismen für Smart Devices. Dabei lassen sich zwei Klassen unterscheiden, für die im Folgenden jeweils ein repräsentativer Vertreter betrachtet wird. Ein Ansatz besteht darin, Datenschutzkomponenten in Anwendungen einzubauen, die dafür Sorge tragen, dass sich eine Anwendung selbst überwacht und Datenschutzeinstellungen einhält. Diese Komponenten können beim Download einer Anwendung automatisch in ihren Bytecode eingebracht werden [Xu12]. *AppGuard* [Ba13] verfolgt diesen Ansatz. Nutzer können damit festlegen, auf welche Daten eine Anwendung Zugriff haben soll und ob die Daten verfremdet weitergegeben werden sollen. Diese Zugriffsregeln können an einen Kontext gebunden werden, in dem sie gültig sind. Der andere Ansatz besteht darin, die Datenschutzkomponenten in das Betriebssystem der Smart Devices einzubringen. *Apex* [Na10] verfolgt diesen Ansatz. Dabei können Zugriffsregeln mit Restriktionen verknüpft werden, die festlegen, wie oft ein bestimmtes Datum abgerufen werden darf.

Diese Ansätze haben jedoch drei entscheidende Probleme: a) Da die Berechtigungen auf Sensoren abgebildet werden, sind diese Systeme für den Industrie 4.0 Anwendungsfall zu restriktiv. Ihre Zugriffsregeln sind somit zu grobgranular, wodurch umfassende Analysen verhindert werden. b) Sie sind nur für einzelne Smart Devices ausgelegt und zielen nicht auf

verteilte Infrastrukturen ab. In der Industrie 4.0 werden hingegen Daten von vielen verschiedenen Geräten gesammelt und verarbeitet. c) Viele dieser Ansätze verstoßen gegen geltendes Recht, z. B. verstößt die Manipulation von Bytecode gegen das Urheberrecht [Al17a].

Datenschutzmechanismen für Datenstromsysteme. Für viele Datenstromsysteme gibt es Erweiterungen, die es ermöglichen, dass bestimmte Attribute eines Datums nur für berechtigte Verarbeitungseinheiten sichtbar sind (z. B. *Borealis* [LM06]). Da dies sehr restriktive Zugriffsregeln zur Folge hat, führt *ACStream* [Ca09] kontextbasierte Zugriffsregeln ein – allerdings ebenfalls ausschließlich auf Attributebene. Wang et al. untersuchen daher im Bereich des *Complex Event Processings*, wie sich eine Zugriffskontrolle auf Ereignisebene (d. h., unter Berücksichtigung von Attributsequenzen) auf die Datenqualität auswirkt [Wa13]. Dies erlaubt wesentlich feingranularere Zugriffsregeln, allerdings geht zu viel Information verloren, wenn bestimmte Ereignisse komplett fallengelassen werden. *PrivApprox* [Qu17] verfolgt daher einen Ansatz, der auf *Differential Privacy* basiert. Das bedeutet, dass detaillierte Daten von verschiedenen Personen analysiert werden, die Ergebnisse jedoch keine Informationen über individuelle Personen preisgeben. Dieses Vorgehen ist allerdings für den Industrie 4.0 Anwendungsfall nicht geeignet, da hier beispielsweise die Aktivitäten einzelnen Werkern zuordenbar sein müssen. Ein weiteres Problem, das all diesen Ansätzen innewohnt ist, dass die Nutzer alle Kontrolle über ihre Daten verlieren, nachdem diese an das Datenstromsystem weitergegeben wurden, und sie blind auf die Einhaltung der Datenschutzrichtlinien vertrauen müssen. Darüber hinaus steigt die Komplexität dieser Ansätze mit der Menge der zu verarbeitenden Daten. Um daher eine Verarbeitung in Echtzeit zu ermöglichen, sollte das Datenvolumen frühzeitig reduziert werden, d. h., Daten, die nicht verarbeitet werden dürfen, sollten bereits auf den Smart Devices herausgefiltert werden.

5 THOR

Wie die Diskussion der verwandten Arbeiten zeigt, muss ein Datenschutzmechanismus für die Industrie 4.0 insbesondere zwei Eigenschaften besitzen. Zum einen muss der Mechanismus feingranulare Berechtigungen unterstützen, damit diese bedarfs- und situationsgerecht vergeben werden können. Sind die Berechtigungen nicht ausreichend feingranular, so ist der Mechanismus zwangsläufig zu restriktiv, worunter die Datenqualität und -vollständigkeit leidet. Zum anderen reicht ein Datenschutzmechanismus für Smart Devices oder für Datenstromsysteme nicht aus. Vielmehr wird eine Kombination die beiden Typen benötigt. Auf den Smart Devices werden die Daten dabei vorgefiltert, damit besonders schützenswerte Daten gar nicht in das Datenstromsystem gelangen und sich dort das Datenvolumen reduziert. Dadurch können im Datenstromsystem komplexe Techniken eingesetzt werden, die den Datenschutz gewährleisten, aber dennoch die Datenqualität berücksichtigen. Diese beiden Mechanismen müssen enge verzahnt sein, um diese Synergieeffekte zu erzielen. Abb. 2 zeigt, wie diese Kombination in THOR realisiert ist. Dabei kommen die **PMP** (**P**rivacy **M**anagement **P**latform) und **PATRON** (**P**riv**a**cy in S**tr**eam Pr**o**cessi**n**g) zum Einsatz. Diese beiden Datenschutzsysteme werden im Folgenden kurz vorgestellt, bevor detailliert wird, wie die Koppelung, Konfiguration und Verifikation der beiden Systeme erfolgt.

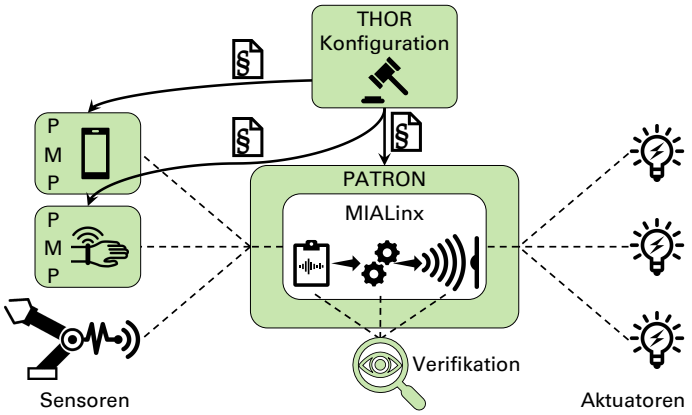

Abb. 2: Datenschutzkonzepte für die *Industrie 4.0* (*THOR*-Komponenten sind in grün dargestellt).

Datenschutz für Smart Devices. Die PMP [SM13; SM14] ist ein Datenschutzsystem für Smart Devices. Bei der PMP kann der Nutzer bestimmte Funktionen einer Anwendung auf Wunsch deaktivieren, wenn er diese nicht benötigt, um so die Datenzugriffe zu reduzieren. Die PMP bietet hierbei fünf wesentliche Eigenschaften: **Berechtigungsänderungen zur Laufzeit** erlauben die Datenschutzeinstellungen jederzeit anzupassen (z. B. um weitere Funktionen einer Anwendung zu aktivieren, wofür diese auf mehr private Daten zugreifen muss). So kann die PMP die Datenweitergabe bedarfsgerecht regulieren. Dies wird dadurch verstärkt, dass sie mit **kontextsensitiven Berechtigungen** arbeitet, d. h., jede Berechtigung ist an einen Aktivierungskontext geknüpft. Im Gegensatz zu AppGuard ist der Kontext allerdings nicht auf spatio-temporale Daten beschränkt, sondern jede verfügbare Datenquelle wird unterstützt. Ein Werker könnte beispielsweise Informationen zu seinen Aktivitäten nur dann weitergeben, wenn er eine neue Tätigkeit ausführt, damit analysiert werden kann, ob er diese korrekt ausführt. Dieses Beispiel zeigt, dass der Kontext bei der PMP wesentlich komplexer sein kann, als nur reine Zeit- und Ortsbezüge. Darüber hinaus ermöglicht die PMP eine **Verfremdung der Daten**, bevor diese einer Anwendung zur Verfügung gestellt werden. Die Verfremdung erfolgt abhängig von der jeweiligen Datenquelle, da sich nicht jede Verfremdungsmethode für jede Datenart eignet. Auf diese Weise ist es möglich, Anwendungen Zugriff auf private Daten zu geben, die Genauigkeit dieser Daten allerdings so weit zu verringern, dass diese keine Bedrohung für die Privatheit des Nutzers darstellen. Da im IoT stetig Smart Devices mit neuartigen Sensoren auf den Markt kommen, ist es für ein Datenschutzsystem sehr wichtig, **erweiterbar** zu sein. In der PMP ist dies dadurch realisiert, dass zur Laufzeit sogenannte *Resources* installiert werden können. Eine Resource ist eine Software-Komponente, die den Zugriff auf eine Datenquelle kontrolliert. Im IoT besitzt ein Nutzer in der Regel mehrere Smart Devices. Daher sollte ein Datenschutzsystem einen Mechanismus zum Austausch von Datenschutzeinstellungen haben. In der PMP ist dies mittels **Presets** realisiert. Darüber können Einstellungen exportiert, an andere Smart Devices gesandt und dort importiert werden.

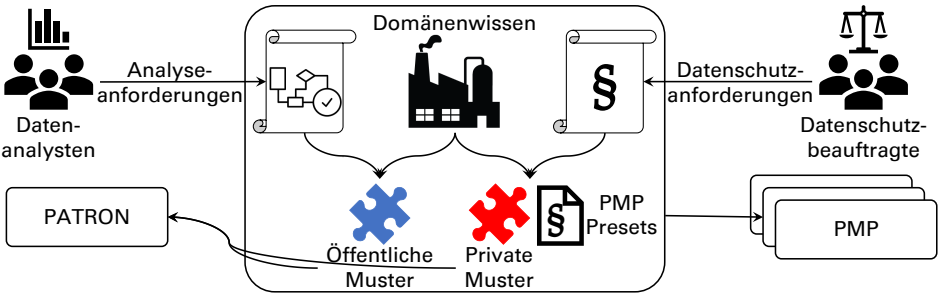

Abb. 3: Spezifikation der Datenschutzeinstellungen in *THOR*.

Aufgrund dieser Eigenschaften kann sich die PMP leicht auf veränderte Bedingungen anpassen. Daher eignet sie sich besonders gut für eine dynamische Umgebung wie die Smart Factory. Die Presets bieten der IT-Abteilung darüber hinaus eine einfache Möglichkeit die Datenschutzeinstellungen auf den Smart Devices der Werker zentral zu managen.

Datenschutz für Datenstromsysteme. PATRON [St18b] ist ein Datenschutzsystem für Datenstromsysteme. Die Datenstromanwendungen werden dabei in PATRON ausgeführt, d. h., PATRON kann alle eingehenden und ausgehenden Datenströme kontrollieren. Im Gegensatz zu den existierenden Arbeiten auf diesem Gebiet, geht PATRON nicht attributbasiert sondern musterbasiert vor. Ein Nutzer definiert **private Datenmuster**, die von PATRON zu schützen sind. Ein Muster ist dabei eine Abfolge von einzelnen Attributen bzw. Sensorwerten. Zum Schutz dieser Muster stehen PATRON dabei mehrere Techniken zur Verfügung, z. B. Umordnung, Unterdrückung oder Verfälschung der einzelnen Attribute. Anwendungen hingegen spezifizieren **öffentliche Datenmuster**[3], die diese im Rahmen der Verarbeitung erkennen können müssen. PATRON versucht daraufhin den bestmöglichen Kompromiss zu finden, so dass keine privaten Muster der Anwendung gegenüber offengelegt werden, aber möglichst viele öffentliche Muster von dieser erkannt werden. Hierfür verfügt PATRON über eine Qualitätsmetrik, die jede mögliche Konfiguration, d. h., jede Verschleierungstechnik bewertet. Dabei wird berücksichtigt, wie viele private Muster offengelegt werden, wie viele öffentliche Muster nicht erkannt werden und wie viele öffentliche Muster falsch erkannt werden, wobei jeder der drei Faktoren situationsabhängig gewichtet werden kann. Die Konfiguration mit dem besten Qualitätswert wird schließlich angewandt.

Da PATRON auf Datenmustern arbeitet und die beschriebene Qualitätsmetrik verwendet, kann Anwendungen eine wesentlich höhere Datenqualität angeboten werden, als bei vergleichbaren Ansätzen. Die angewandten Restriktionen sind wesentlich feingranularer, da nicht der Zugriff auf bestimmte Attribute untersagt wird, sondern stets die zeitliche Abfolge berücksichtigt wird. Außerdem werden durch öffentliche Muster auch die Anforderungen der Anwendungen miteinbezogen, wodurch trotz Einhaltung des Datenschutzes der zuverlässige Einsatz eines Systems wie MIALinx überhaupt erst ermöglicht wird.

[3] „Öffentliche Datenmuster“ sind nicht frei zugänglich, sondern zur rechtskonformen Datenauswertung freigegeben.

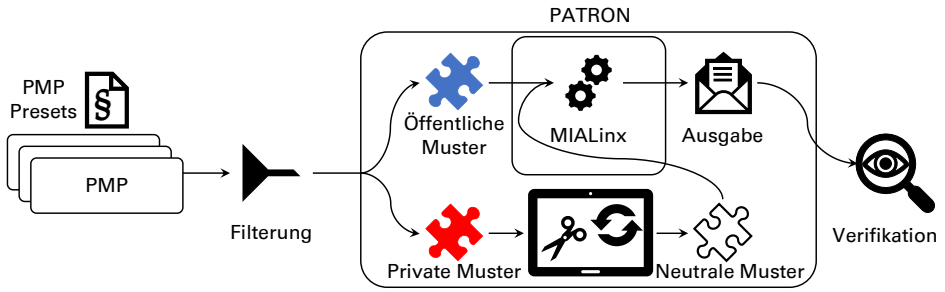

Abb. 4: Verifikation der Datenschutzeinstellungen in *THOR*.

Kombination der Datenschutzsysteme. Die PMP und PATRON bieten auf ihrem jeweiligen Gebiet einen umfassenden Datenschutz. THOR führt daher diese beiden Systeme zusammen, um ein holistisches Datenschutzsystem für die Industrie 4.0 zu erhalten. Abb. 3 zeigt, wie in THOR die Spezifikation der Datenschutzeinstellungen erfolgt. Analog zu PATRON wird dabei auf spezifisches Domänenwissen zurückgegriffen, z. B. darüber, welche Rückschlüsse ein bestimmter Sensorwert respektive eine Abfolge von Sensorwerten erlauben. Datenanalysten geben ihre Analyseziele an, woraus semi-automatisch öffentliche Muster für die jeweilige Domäne abgeleitet werden können. Die Datenschutzbeauftragten beschreiben die Schutzziele, woraus die privaten Muster gebildet werden. Zusätzlich bestimmt THOR, welche Daten auf den Smart Devices erhoben werden, die nicht oder nur in aggregierter Form für die Analysen benötigt werden. Hieraus entstehen die Presets für die PMP. Die PATRON Qualitätsmetrik muss hierbei versuchen möglichst viele Daten frühzeitig herauszufiltern und nur ein Mindestmaß an das Datenstromsystem weiterleiten. Die Konfigurationen werden daraufhin an die beiden Datenschutzsysteme geschickt. Wie solche Konfigurationen in einer heterogenen Infrastruktur ausgebracht werden können, ist in [Al17b] beschrieben.

Abb. 4 zeigt die Arbeitsweise von THOR nach der Konfiguration. Zur Laufzeit sichert die PMP sämtliche Datenquellen ab. Nur Daten, die für die Analyse unerlässlich sind, werden zur weiteren Verarbeitung freigegeben (z. B. durch MIALinx). Der Verarbeitungsmechanismus ist allerdings vollständig in PATRON eingebettet. Somit kennt PATRON alle eingehenden Datenströme und kann diese auf die definierten Muster hin untersuchen. Wird ein öffentliches Muster entdeckt, so wird dieses unverändert an MIALinx weitergegeben. Wird hingegen ein privates Muster erkannt, so wird dieses in ein neutrales, d. h., nicht-öffentliches Muster überführt. Sämtliche Ausgaben von MIALinx können von PATRON mitgelesen werden. Datenschutzbeauftragte und Analysten können dadurch verifizieren, ob private Daten preisgegeben oder relevante Ereignisse nicht erkannt wurden und anschließend zusätzliche private respektive öffentliche Muster definieren. Durch dieses Zusammenspiel der PMP und PATRON wird THOR zu einem Datenschutzsystem für die Industrie 4.0, das den Schutz personenbezogener Daten sicherstellt, dabei aber auch die für Analysen erforderliche Datenqualität angemessen berücksichtigt. Im Folgenden muss nun untersucht werden, ob dadurch alle Datenschutzanforderungen (siehe Abschnitt 3) erfüllt werden.

6 Evaluation

THOR setzt auf allen Geräten, die personenbezogene Daten erfassen die PMP ein. Damit verfügt es über ein Datenschutzsystem speziell für Smart Devices ($\mathbf{R_1}$). Über den zentralen Konfigurationsmechanismus in THOR können sowohl die Analyseziele als auch die Datenschutzziele spezifiziert werden. Aus diesen beiden Spezifikationen werden anschließend Konfigurationen für die PMP und PATRON gebildet und auf alle Smart Devices und das Datenstromsystem verteilt (siehe Abb. 3). Damit besitzt THOR eine einheitliche Konfiguration ($\mathbf{R_2}$). Die PMP verwendet kontextsensitive Berechtigungen, dadurch wird ein kontextabhängiger Datenschutz ermöglicht. Auch PATRON kann über die Muster den Kontext von Daten berücksichtigen ($\mathbf{R_3}$). Mit PATRON kommt in THOR ein Datenschutzsystem für Datenstromsysteme und -anwendungen, wie beispielsweise MIALinx zum Einsatz ($\mathbf{R_4}$). Dabei führt PATRON öffentliche und private Muster ein ($\mathbf{R_5}$). Über eine Qualitätsmetrik kann gezielt dafür gesorgt werden, dass die Datenqualität berücksichtigt wird (öffentliche Muster werden korrekt erkannt) während der Datenschutz aufrechterhalten wird (private Muster werden verborgen) ($\mathbf{R_6}$). Die Ausgabe des Datenstromsystems kann von THOR eingesehen werden. Dadurch können Nutzer nachvollziehen, ob die Konfiguration zu restriktiv ist (öffentliche Muster fehlen) oder zu viele private Daten preisgegeben werden (private Muster fehlen), und entsprechende Anpassungen vornehmen ($\mathbf{R_7}$).

Somit werden von THOR alle Datenschutzanforderungen an ein Datenschutzkonzept für die Industrie 4.0 erfüllt.

7 Zusammenfassung

Bedingt durch den Aufschwung des IoTs werden über immer mehr Aspekte unseres Alltags Daten erfasst und analysiert. In kaum einer Domäne haben diese Daten einen vergleichbaren ökologischen Wert, wie in der Fertigung. Erst die umfassende Bestückung cyber-physischer Systeme mit Sensoren ermöglicht die Industrie 4.0. Zunehmend werden dabei auch Daten über Werker gesammelt (z. B. mittels Smart Watches). Bei diesen personenbezogenen Daten muss allerdings die DSGVO beachtet werden. Die heutigen technischen Datenschutzmechanismen sind für diesen Anwendungsfall allerdings zu restriktiv und decken nur Teilaspekte ab. Daher bedarf es an neuen, innovativen Datenschutzkonzepten.

Im Rahmen dieser Arbeit führen wir mit THOR ein solches Konzept speziell für die Industrie 4.0 ein. In THOR wird mit der PMP und PATRON ein Datenschutzsystem für Smart Devices mit einem Datenschutzsystem für Datenstromsysteme verknüpft. Mit der PMP werden die erhobenen Daten direkt auf den Smart Devices vorgefiltert und nur prozessrelevante Daten werden zur weiteren Verarbeitung durch ein Datenstromsystem freigegeben. Dort greift PATRON und sorgt dafür, dass keine als privat deklarierten Muster (d. h., Abfolgen von Sensorwerten) offengelegt werden können. Hierfür stehen PATRON unterschiedliche Techniken (z. B. Umordnung oder Unterdrückung) zur Verfügung. Da zwischen der PMP und PATRON eine enge Kopplung besteht, können die Datenschutzaufgaben so aufgeteilt

werden, dass die Datenqualität berücksichtigt wird, d. h., die für die Produktion benötigten Analysen bleiben davon weitestgehend unberührt. Die Evaluation des Gesamtsystems zeigt, dass THOR alle Datenschutzanforderungen, die sich im Umfeld der Industrie 4.0 ergeben, erfüllt und sich damit die DSGVO in der Smart Factory technischen realisieren lässt.

Danksagung

Wir danken der Baden-Württemberg Stiftung für die Förderung der in diesem Artikel vorgestellten Forschungsarbeiten. Bei MIALinx und PARTON handelt es sich um Forschungsaufträge, die aus Mitteln der Baden-Württemberg Stiftung finanziert werden.

Literatur

[AA16] Albaghli, R.; Anderson, K. M.: A Vision for Heart Rate Health Through Wearables. In: UbiComp ’16. 2016.

[Al17a] Alpers, S. et al.: Herausforderungen bei der Entwicklung von Anwendungen zum Selbstdatenschutz. In: 47. GI-Jahrestagung. 2017.

[Al17b] Alpers, S. et al.: PRIVACY-AVARE: An Approach to Manage and Distribute Privacy Settings. In: ICCC ’17. 2017.

[Ba13] Backes, M. et al.: AppGuard — Fine-Grained Policy Enforcement for Untrusted Android Applications. In: DPM ’13. 2013.

[Ca09] Cao, J. et al.: ACStream: Enforcing Access Control over Data Streams. In: ICDE ’09. 2009.

[Co11] Conti, M. et al.: Mind How You Answer Me!: Transparently Authenticating the User of a Smartphone when Answering or Placing a Call. In: ASIACCS ’11. 2011.

[EU16] Das Europäische Parlament und der Rat der Europäischen Union: Verordnung (EU) 2016/679 des Europäischen Parlaments und der Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung), Amtsblatt, Europäischen Union, 2016.

[Ha15] Hashem, I. A. T. et al.: The rise of “big data” on cloud computing: Review and open research issues. Information Systems 47/1, S. 98–115, 2015.

[Hs14] Hsu, H.-H. et al.: Smartphone Indoor Localization with Accelerometer and Gyroscope. In: NBiS ’14. 2014.

[Ka17] Kassner, L. et al.: The Social Factory: Connecting People, Machines and Data in Manufacturing for Context-Aware Exception Escalation. In: HICSS ’17. 2017.

[Kh17] Khan, M. et al.: Big data challenges and opportunities in the hype of Industry 4.0. In: ICC ’17. 2017.

[Ki17] Kim, G. et al.: Mobile Security Solution for Sensitive Data Leakage Prevention. In: ICCBN ’17. 2017.

[LM06] Lindner, W.; Meier, J.: Securing the Borealis Data Stream Engine. In: IDEAS ’06. 2006.

[Mi13] Middleton, P. et al.: Forecast: The Internet of Things, Worldwide, Report, Gartner, Inc., 2013.

[Na10] Nauman, M. et al.: Apex: Extending Android Permission Model and Enforcement with User-defined Runtime Constraints. In: ASIACCS ’10. 2010.

[OB16] Oluwatimi, O.; Bertino, E.: An Application Restriction System for Bring-Your-Own-Device Scenarios. In: SACMAT ’16. 2016.

[Qu17] Quoc, D. L. et al.: PrivApprox: Privacy-Preserving Stream Analytics. In: ATC ’17. 2017.

[Se17] Seo, J. et al.: An Analysis of Economic Impact on IoT under GDPR. In: ICTC ’17. 2017.

[Sh17] Shahmohammadi, F. et al.: Smartwatch Based Activity Recognition Using Active Learning. In: CHASE ’17. 2017.

[SM13] Stach, C.; Mitschang, B.: Privacy Management for Mobile Platforms – A Review of Concepts and Approaches. In: MDM ’13. 2013.

[SM14] Stach, C.; Mitschang, B.: Design and Implementation of the Privacy Management Platform. In: MDM ’14. 2014.

[St18a] Stach, C.: Big Brother is Smart Watching You: Privacy Concerns about Health and Fitness Applications. In: ICISSP ’18. 2018.

[St18b] Stach, C. et al.: How a Pattern-based Privacy System Contributes to Improve Context Recognition. In: CoMoRea ’18. 2018.

[UA16] Urbach, N.; Ahlemann, F.: Der Wissensarbeitsplatz der Zukunft: Trends, Herausforderungen und Implikationen für das strategische IT-Management. HMD Praxis der Wirtschaftsinformatik 53/1, S. 16–28, 2016.

[Wa13] Wang, D. et al.: Utility-maximizing Event Stream Suppression. In: SIGMOD ’13. 2013.

[Wi16] Wieland, M. et al.: Towards a Rule-based Manufacturing Integration Assistant. Procedia CIRP 57/1, S. 213–218, 2016.

[Wi17] Wieland, M. et al.: Rule-Based Integration of Smart Services Using the Manufacturing Service Bus. In: UIC ’17. 2017.

[Xu12] Xu, R. et al.: Aurasium: Practical Policy Enforcement for Android Applications. In: Security ’12. 2012.