Institut für Architektur von Anwendungssystemen

Universität Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Masterarbeit

Entwicklung eines Decision Support Systems für die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen

Johannes Hepp

Studiengang:	Informatik
Prüfer/in:	Prof. Dr. Dr. h.c. Frank Leymann
Betreuer/in:	Dr. Uwe Breitenbücher, Dr. Tobias Weber
Beginn am:	6. April 2022
Beendet am:	20. Oktober 2022

Kurzfassung

Seit dem Aufkommen der Cloud hat sich die Art Anwendungen zu betreiben stark verändert: Die Cloud ist dezentral und einfach skalierbar. Durch das Prinzip *pay-per-use* ist der Betrieb von Komponenten in Cloud-Umgebungen für viele Firmen attraktiv. Allerdings entstehen bei der Auslagerung von Komponenten in Cloud-Umgebungen immer wieder Probleme und Schwierigkeiten, besonders bei Anwendungen, die sicherheitskritische Daten verarbeiten, die bisher in firmeninternen Infrastrukturen betrieben wurden.

In den letzten Jahren wurden für diese Probleme zahlreiche Ansätze und Technologien entwickelt, allerdings ist der Zugang zu diesem Wissen erschwert, da es in unterschiedlichsten Formen dokumentiert ist, wie beispielsweise wissenschaftliche Literatur, White Papers oder Blog-Posts, und die passende Lösung von vielen Variablen abhängt.

Aus diesem Grund beschäftigt sich diese Arbeit damit, dieses Wissen mit Hilfe der Entwicklung eines Decision Support Systems einfach zugänglich zu machen. Durch die Befragung von Experten werden verschiedene Arten von sicherheitskritischen Daten festgelegt, sowie Anforderungen für das Decision Support System gesammelt. Für die verschiedenen Arten von sicherheitskritischen Daten werden existierende Best Practices und Patterns für den Betrieb von Anwendungen, die diese Daten verarbeiten, recherchiert und eigene Patterns erstellt.

Bei den Experteninterviews wurde klar, dass ein System das alle wichtigen Informationen zusammenfasst bisher nicht existiert und eine zentrale Lösung einen großen Mehrwert in der täglichen Arbeit bieten würde. Desweiteren konnten mit Hilfe einer Multivocal Literature Review mehrere Pattern-Sprachen ermittelt werden, die Informationen über den Betrieb von sicherheitskritischen Daten in der Cloud enthalten und drei eigene Pattern erstellt werden, die die rechtlichen Problematiken mit sicherheitskritischen Daten abbilden. Der Prototyp, der im Rahmen dieser Arbeit entstanden ist, macht dieses Wissen auf übersichtliche und einfache Weise zugänglich.

Zusammengefasst bietet diese Masterarbeit Informationen für die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen mit einer Auflistung und Erstellung relevanter Pattern-Sprachen und einem einfachen Zugang durch einen Prototypen. Der Schutz dieser Daten ist besonders relevant, da ein Verlust große Imageschäden und finanzielle Belastungen verursachen kann.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Abkürzungsverzeichnis

AWS Amazon Web Services. 27

DSGVO Datenschutz Grundverordnung. 13

EU Europäische Union. 13, 20, 26, 38, 39, 45, 47, 48, 50, 51, 52, 53, 63, 64

EUGH Europäische Gerichtshof. 20

GC Google Cloud. 27

GeschGehG Gesetz zum Schutz von Geschäftsgeheimnissen. 38

GL Gray Literature. 19

IT Informationstechnologie. 22

MLR Multivocal Literature Review. 14

MSA Microsoft Azure. 28

SIT Schwarz IT KG. 20

SLR Systematic Literature Review. 19

1 Einleitung

Die Cloud hat den Betrieb von Anwendungen stark verändert. Sie ist *die* dezentrale und skalierbare Hostinglösung für viele Unternehmen geworden. Doch bei der Auslagerung von Anwendungen in die Cloud kommt es immer wieder zu Problemen, gerade wenn diese Anwendungen sicherheitskritische Daten verarbeiten. Diese Aufgabe wird durch besonders strenge Datenschutzgesetze für personenbezogene Daten innerhalb der Europäischen Union (EU) erschwert. In der Vergangenheit wurden für diese Probleme bereits zahlreiche Lösungen entwickelt. Der Zugang zu diesem Wissen ist allerdings stark erschwert, da es in den unterschiedlichsten Formen von unterschiedlichen Institutionen festgehalten wurde, wie wissenschaftlichen Arbeiten, White-Papers, Blog-Posts oder Dokumentationen.

Die Datenschutz Grundverordnung (DSGVO) [Sch19] regelt die Verarbeitung von personenbezogenen bzw. personenbeziehbaren Daten innerhalb der EU. Neben verschiedenen Regeln, wie beispielsweise zum Umfang der Erfassung dieser Daten, Aufbewahrungsfristen oder Möglichkeit der permanenten Löschung, regelt die DSGVO auch die Pflichten bei der Aufbewahrung und Verarbeitung dieser Daten. Grundlegend müssen die Daten mit Sicherheitsmaßnahmen ausreichend durch den Eigentümer geschützt werden. Der Kontext der Cloud macht diese Thematik allerdings komplexer, denn die DSGVO regelt auch die Übermittlung von personenbezogenen Daten an Drittländer oder internationale Organisationen. Da es sich bei vielen Cloud-Anbietern um internationale Unternehmen handelt, sind hier diese Regelungen zu beachten. Personenbezogene Daten dürfen nur dann an Länder oder Organisationen außerhalb der EU zur Verarbeitung übermittelt werden, wenn diese ein angemessenes Schutzniveau bieten. Ein angemessenes Schutzniveau wird unter anderem so definiert, dass das Datenschutzniveau im Drittland mindestens dem der EU entsprechen muss. Welche Auswirkungen diese Regelungen auf die Wahl des Anbieters bzw. die Art der Verarbeitung hat wird in dieser Arbeit untersucht.

Das Ziel dieser Arbeit ist das Wissen über die aktuelle Rechtslage und die dazu passenden technischen Lösungen zusammenzutragen und einfach zugänglich zu machen. Dafür soll ermittelt werden, welche Arten von sicherheitskritischen Daten in der Praxis existieren und verschiedene Patterns und Pattern-Sprachen gefunden und gegebenenfalls erstellt werden, die bewährte Lösungen für die verschiedenen Probleme anbieten, die beim Betrieb von diesen verschiedenen Arten von sicherheitskritischen Daten in der Cloud auftreten können. Diese Pattern-Sprachen sollen dann mithilfe eines Decision Support Systems übersichtlich und einfach für jeden zugänglich gemacht werden.

Für die Ermittlung der verschiedenen Arten von sicherheitskritischen Daten in der Praxis und den Problemstellungen die diese mit sich bringen werden Experteninterviews durchgeführt. Die Interviews dienen zusätzlich zur Anforderungsanalyse des Decision Support Systems. Da dieses sich direkt an Entwickler, Architekten und Forschende richtet, sind die Experten auch die spätere

Nutzergruppe. Für die Recherche der verschiedenen Pattern und Pattern-Sprachen wurde eine Multivocal Literature Review (MLR) durchgeführt, um somit die Daten für das Decision Support System zu erhalten.

Die Arbeit ist wie folgt strukturiert:

- **Kapitel 2 - Grundlagen**

  In diesem Kapitel werden verschiedenste Konzepte vorgestellt, die für das spätere Verständnis der Arbeit benötigt werden.

- **Kapitel 3 - Multivocal Literature Review**

  Der gesamte MLR-Prozess wird vorgestellt, besonderes Augenmerk wird auf die Motivation des MLR gelegt. Dort sind ebenfalls die Experteninterviews dokumentiert.

- **Kapitel 4 - Pattern-Sprachen für die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen**

  In diesem Kapitel werden die verschiedenen Pattern-Sprachen vorgestellt, die im Rahmen der MLR gefunden wurden.

- **Kapitel 5 - Data Protection Pattern Language**

  Hier wird die Pattern-Sprache vorgestellt die im Rahmen dieser Arbeit erstellt wurde.

- **Kapitel 6 - Decision Support System**

  In diesem Kapitel wird der Prototyp vorgestellt, der im Rahmen dieser Arbeit entstanden ist. Die Anforderungen für diesen Prototypen wurden dabei in Kapitel 3 mithilfe des Experteninterviews ermittelt.

- **Kapitel 7 - Zusammenfassung und Ausblick**

  Dieses Kapitel enthält die Zusammenfassung dieser Arbeit sowie einen Ausblick für zukünftige Arbeiten.

2 Grundlagen

In den folgenden Abschnitten werden grundlegende Konzepte vorgestellt, die für das Verständis der Arbeit notwendig sind. In Abschnitt 2.1 wird definiert, was Best Practices sind. Danach wird in Abschnitt 2.2 das Konzept von Patterns erklärt und in Abschnitt 2.3 das Konzept von Pattern-Sprachen vorgestellt.

2.1 Best Practices

Als Best Practice bezeichnet man bewährte beziehungsweise optimale Verfahren für ein gegebenes Problem. Diese Lösungen können in unterschiedlichsten Bereichen, wie Wirtschaft, Bildung, Gesundheit oder Informatik vorkommen. Best Practices werden meist von Unternehmen oder Gesetzgebern definiert und bestimmen das allgemein sinnvollste Vorgehen. Anders als ein offizieller Standard oder Normen können sich Best Practices auch über die Zeit ändern oder sogar obsolet werden. Sie dienen als generelles Framework für verschiedenste Probleme und bieten für Unternehmen die Möglichkeit bestmöglich mit Problemen umzugehen. Gern werden Best Practices auch verwendet, um ein firmenweit einheitliches Vorgehen zu gewährleisten. [Inv22]

2.2 Konzept der Patterns

Bei Pattern handelt es sich um ein Konzept, das 1977 das erste Mal von Christopher Alexander und Mitwirkende in *A pattern language: towns, buildings, custruction* [Ale77] eingeführt wurde. Bei Pattern handelt es sich, um eine Dokumentation von Techniken, um eine Klasse von Problemen zu lösen. Pattern bilden dabei nicht eine exakte Lösungsanleitung für ein bestimmtes Problem, sondern beschreiben Lösungswege beziehungsweise Lösungsansätze, um wiederkehrende Probleme zu lösen. Dabei dokumentiert ein Pattern auch alle Dinge, die zu der entsprechenden Lösung geführt haben und warum es sich bei der gegebenen Lösung um die bestmögliche handelt. Pattern sind in folgende Struktur unterteilt:

- Name
  - Beschreibt das Pattern und sagt direkt aus, welches Problem das Pattern löst.
- Icon
  - Beschreibt visuell das Pattern und liefert einen Wiedererkennungswert. Durch ein Icon lässt sich das Pattern ebenfalls sehr einfach in Skizzen anderer Patterns wiederverwenden.
- Problem
  - Hier wird beschrieben, mit welchen Problemen man konfrontiert ist.

- Kontext
  - Es werden die Umstände für das Problem beschrieben, also wie es zu diesem Problem kam. Was sind die Vorbedingungen, um das Pattern anwenden zu können und welche Seiteneffekte bei der Anwendung dieses Patterns auftreten.
- Motivation
  - In diesem Bereich wird beschrieben, warum das Problem schwierig zu lösen ist und warum bestehende Lösungen oder Alternativen nicht funktionieren.
- Lösung
  - Es wird die Lösung für das Problem aufgezeigt.
- Skizze
  - Stellt die Lösung für das Problem visuell dar und hilft damit die beschriebene Lösung nachzuvollziehen.
- Ergebnisse
  - Hier wird beschrieben, wie die Lösung angewendet werden kann und wie die Probleme, die in Motivation beschrieben wurden gelöst werden.
- Ausblick
  - Hier wird der Zusammenhang zu anderen Patterns beschrieben
- Variationen
  - In diesem Bereich werden alternative Patterns beschrieben, falls vorhanden.
- Beispiele
  - Hier werden verschiedene konkrete Anwendungsbeispiele der Lösung präsentiert.

2.3 Konzept der Pattern-Sprachen

Bei einer Pattern-Sprache handelt es sich um einen Verbund aus Patterns, die zuvor in 2.2 beschrieben wurden. Das Konzept einer Pattern-Sprache wurde zuerst in *A pattern language: towns, buildings, custruction* [Ale77] im Kontext der Architektur vorgestellt um bewiesene Lösungen für wiederkehrende Probleme zu bieten. Eine Pattern-Sprache, wie sie damals eingeführt wurde, verbindet Patterns logisch miteinander. Dadurch lassen sich Zusammenhänge zwischen diesen erkennen und darstellen. Die einfachste Darstellung einer Pattern-Sprache ist dabei ein ungerichteter Graph. Eines der Probleme der Darstellung als ungerichteter Graph ist, dass jedes Pattern einzeln betrachtet werden muss, um den Zusammenhang zu erkennen, der in den verschiedenen Bereichen eines Patterns beschrieben wird, was die Verwendung der Zusammenhänge sehr umständlich gestaltet. Aus diesem Grund wird in dieser Arbeit die Darstellung einer Pattern-Sprache verwendet, die in *The Nature of Pattern Languages* [FBL18] vorgestellt wurde. Dort wurde der Graph einer Pattern-Sprache als 6-Tupel definiert.

Bei N handelt es sich um die Menge von Patterns und bei E um die Menge von Kanten zwischen den Patterns. Bei $\mathfrak{D}$ handelt es sich um eine Menge von Typen die einer Kante zugewiesen werden können. Mit Hilfe dieser Typen wird die Beziehung zwischen den Patterns genauer beschrieben. $\mathcal{W}$ sind die Gewichte. Diese können den Kanten zugewiesen werden und können bei Typen eingesetzt werden, die diese für spezifischere Beziehungen benötigen. Daraus wird der folgende gewichtete Graph $\mathcal{G}$ definiert:

$\mathcal{G} = (N, E, \mathcal{W}, \mathfrak{D}, \alpha, \beta)$

mit

1. N ist eine Menge von Patterns
2. $card(N) \in \mathbb{N}$
3. $E \subseteq N \times N \times \mathcal{W}$
4. $\mathcal{W} \neq \emptyset$
5. $e \in E$, wenn $\pi_1(e)$ der Startpunkt der Kante e, $\pi_2(e)$ der Endpunkt der Kante e und $\pi_3(e)$ der Typ der der Kante e zugewiesen wurde ist
6. $\forall e \in E : \pi_1(e) \neq \pi_2(e)$
7. $n_1, n_2, ..., n_k \in N$ ist ein Pfad von n_1 zu $n_k :\Leftrightarrow (n_1, n_2), (n_2, n_3), ..., (n_{k-1}, n_k) \in E$
8. Ein Pfad $n_1, n_2, ..., n_k \in N$ ist ein simpler Pfad:$\Leftrightarrow \forall 2 \leq i, j \leq k-1 : n_i \neq n_j$
9. $\forall n_i, n_{i+1}, ..., n_k \in N$ die simple Pfade sind gilt $n_i \neq n_k$
10. $\forall e_i, e_k \in E : \pi_1(e_i) = \pi_1(e_k) \wedge \pi_2(e_i) = \pi_2(e_k) \Rightarrow \pi_3(e_i) \neq \pi_3(e_k)$
11. $\alpha : \mathcal{W} \rightarrow \wp(\mathfrak{D})$
12. $\beta : E \rightarrow \bigcup_{e \in E} \times_{D \in \mathfrak{D}_{\alpha(\pi_3(e))}} D$
13. $\forall e \in E : \beta(e) \in \times_{D \in \mathfrak{D}_{\alpha(\pi_3(e))}} D$

wo α eine Map ist, die Teilmengen aller Domänen ein Gewicht zuweist und β eine Map ist, die Kanten typspezifische Beschreibungen zuweist ■. [FBL18]

3 Multivocal Literature Review

Für die Durchführung der MLR wurde das in Garousi's Artikel [GFM19] vorgestelle Verfahren verwendet. Im Gegensatz zu einer Systematic Literature Review (SLR), bei der nur wissenschaftliche Arbeiten als Quellen dienen, wird bei einer MLR auch Gray Literature (GL) mit einbezogen. In Abschnitt 3.1 werden die Gründe diskutiert, die für eine MLR sprechen. Um die Forschungsfragen, Such-Strings und Konzepte der MLR zielgerichtet festlegen zu können wurden im Rahmen dieser Arbeit Experteninterviews durchgeführt. Dieser Prozess wird in Abschnitt 3.2 vorgestellt. In Abschnitt 3.3 wird der ausgewählte Suchprozess erklärt und in Abschnitt 3.4 wird vorgestellt, wie die Daten erhoben wurden. In Abschnitt 3.6 werden die Ergebnisse der MLR präsentiert.

3.1 Gründe für die MLR

Garousi [GFM19] hat in seiner Arbeit sechs Kriterien vorgestellt, die für eine MLR gegenüber einer SLR sprechen. Je mehr dieser Kriterien auf das Thema zutreffen, desto notwendiger ist es Gray Literature mit in die Review einzubeziehen.

In Tabelle 3.1 sind die Kriterien aus Garousi's Arbeit aufgelistet und bewertet, welche Kriterien auf das Thema: *Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen* zutreffen. Im Folgenden wird die Entscheidung für jeden dieser Kriterien kurz begründet.

- *Kontextabhängig*

  Durch die vielen Möglichkeiten und die vielen rechtlichen Aspekte für den Betrieb von sicherheitskritischen Daten in der Cloud sind geeignete Lösungen sehr kontextabhängig.

- *Komplexes Ergebnis*

  Ob das Ergebnis der MLR komplex ist, kann zu diesem Zeitpunkt noch nicht eindeutig bestimmt werden. Aus diesem Grund ist dieses Kriterium mit *Nein* bewertet.

- *Fehlender Konsens in überprüfter Literatur*

  Im Kontext der gesetzlichen Regelungen und ihren Auswirkungen auf den Betrieb von personenbezogenen Daten konnte ein Dissens in der überprüften Literatur festgestellt werden. In Selzer's Fachbeitrag [Sel20] heißt es: „[...] dass genehmigte Zertifizierungsverfahren [...] eine geeignete Garantie für die Übermittlung an Drittstaaten darstellen können.“. Wissenschaftliche Dienste des Deutschen Bundestages kommt allerdings in ihrer Ausarbeitung [Die21] zu dem Schluss, dass Garantien allein nicht ausreichend sind, wie hier beschrieben: „In diesem Fall müssen aber zusätzliche Maßnahmen getroffen werden [...]“. Aus diesem Grund ist dieses Kriterium als zutreffend bewertet.

Kriterium	
Kontextabhängig	Ja
Komplexes Ergebnis	Nein
Fehlender Konsens in überprüfter Literatur	Ja
Geringe Evidenz in überprüfter Literatur	Ja
Geringe Qualität der Evidenz in überprüfter Literatur	Nein
Kontext wichtig für die Umsetzung	Ja
Anzahl der *Ja* Antworten	**4**

Tabelle 3.1: Kriterien für eine MLR

- *Geringe Evidenz in überprüfter Literatur*

  2020 hat der Europäische Gerichtshof (EUGH) ein Urteil betreffend der DSGVO gefällt [20], das einige Änderungen des Sachbestandes mit sich bringt. Dadurch, dass sehr viel der überprüften Literatur zu diesem Thema aus 2018 und früher stammt, sind die Informationen daraus nicht sicher verwendbar und die Anzahl der verwertbaren Literatur ist dadurch sehr stark eingeschränkt. Aus diesem Grund gibt es eine geringe Anzahl an Evidenz.

- *Geringe Qualität der Evidenz in überprüfter Literatur*

  Beim Thema Datensicherheit existiert sehr viel Evidenz von wissenschaftlichen Quellen, wodurch hier eine geringe Qualität in der überprüften Literatur nicht festgestellt werden konnte.

- *Kontext wichtig für die Umsetzung*

  Die EU bringt durch die strengen Datenschutzgesetze viele zusätzliche Anforderungen in das Thema Datensicherheit. Durch die vielseitigen Datenverarbeitungsfunktionen der Cloud, ist der Kontext für die Umsetzung innerhalb der EU besonders relevant, denn die Nutzung von personenbezogenen Daten ist stark eingeschränkt und nur besimmte Daten dürfen unter bestimmten Umständen beispielsweise frei verarbeitet werden.

Da insgesamt vier der sechs von Garousi [GFM19] vorgestellten Kriterien auf das Thema zutreffen, empfiehlt es sich GL mit in die Review einzubeziehen.

3.2 Experteninterviews im Rahmen der MLR

Wie Garousi in seiner Arbeit [GFM19] beschrieben hat, ist es eine sehr wichtige Aufgabe die Forschungsfragen einer MLR möglichst zielgerichtet und konkret zu stellen. Denn sie entscheiden, ob das Ergebnis auch einen Mehrwert bietet. Aus diesem Grund wurden im Rahmen dieser Arbeit Interviews mit insgesamt 10 Mitarbeitern der Schwarz IT KG (SIT) und ihrer Dienstleister durchgeführt. Dieser Prozess und dessen Ergebnis wurde in den folgenden Abschnitten festgehalten. Dafür werden zunächst in Abschnitt 3.2.1 die Ziele der Interviews erläutert. In Abschnitt 3.2.2 wird der gewählte Interviewprozess und damit auch die Reihenfolge der Fragen erklärt. Im folgenden Abschnitt 3.2.3 werden die Hintergründe und aus welchen Bereichen und Unternehmen

die verschiedenen Teilnehmer kommen, vorgestellt. Anschließend werden in Abschnitt 3.2.4 die Fragen vorgestellt, die den Experten gestellt werden. Am Ende werden dann in Abschnitt 3.2.5 die Antworten der Teilnehmer präsentiert.

3.2.1 Ziele der Experteninterviews

Eines der Ziele der Interviews ist, die praktische Relevanz der Arbeit zu zeigen. Durch die Interviews soll validiert werden, dass sie auch einen tatsächlichen Mehrwert bietet und in der Zukunft Verwendung findet. Zusätzlich soll herausgefunden werden, welche Arten von sicherheitskritischen Daten in der Praxis existieren und daher in dieser Arbeit betrachtet werden sollen. Das Hauptziel der Interviews ist eine Anforderungsanalyse für das Decision Support System, das im Rahmen dieser Arbeit umgesetzt wird. Aus diesem Grund soll ermittelt werden, welche Werkzeuge und Webseiten bisher von den Experten eingesetzt werden, um Informationen zur Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen zu erhalten und welche Funktionen hier besonders hilfreich sind oder sein könnten.

3.2.2 Prozess des Experteninterviews

Wie im Buch *Case Study Research in Software Engineering: Guidlelines and Examples* [HRRR12] beschrieben, gibt es drei verschiedene Arten Interviews durchzuführen: unstrukturierte, semistrukturierte und strukturierte Interviews. Diese unterscheiden sich im Fokus, Ziel und daraus folgend auch von der Art der Fragen.

Bei unstrukturierten Interviews ist das Ziel möglichst viel Neues zu lernen. Aus diesem Grund werden die Interviews so aufgebaut, dass man bestimme Felder fokussiert und Informationen dafür sammelt. Dadurch werden hier nur offene Fragen verwendet. Das bedeutet, dass keinerlei Antworten vorgegeben werden.

Beim strukturierten Interview ist das Ziel, Zusammenhänge zwischen Themen zu erkennen und Meinungen über diese Themen einzuholen. Aus diesem Grund werden hier nur geschlossene Fragen gestellt. Bei geschlossenen Fragen hat der Interviewpartner nur die Möglichkeit zwischen verschiedenen vorgegebenen Antworten auszuwählen.

Bei semistrukturierten Interviews handelt es sich um eine Mischung zwischen dem unstrukturierten und strukturierten Interviewprozess. Es geht hier darum, sowohl neue Informationen zu sammeln als auch Meinungen zu bestimmten Themen einzuholen. Aus diesem Grund wird hier eine Mischung aus offenen und geschlossenen Fragen verwendet. Dadurch können die Interviewpartner bei bestimmten Fragen sehr offen reden und viele neue Informationen liefern und bei anderen Fragen sehr präzise Meinungen geben, die gut vergleichbar sind. Durch das in Abschnitt 3.2.1 formulierte Forschungsziel eignet sich ein semistrukturiertes Interview, da sowohl viele neue Informationen, als auch Meinungsbilder über bestimmte Themen gesammelt werden sollen.

Für die Planung des Interviews und damit der Reihenfolge der Fragen gibt es drei verschiedene Ansätze, wie von Runeson und Kollegen beschrieben [HRRR12]. Beim Funnel werden zunächst sehr offene Fragen gestellt und später im Interviewprozess geschlossenere Fragen. Bei der Pyramide wird zunächst mit geschlossenen Fragen begonnen, die dann während dem Prozess immer offener werden. Der dritte Ansatz ist das Timeglass. Beim Timeglass wird zunächst mit offenen Fragen

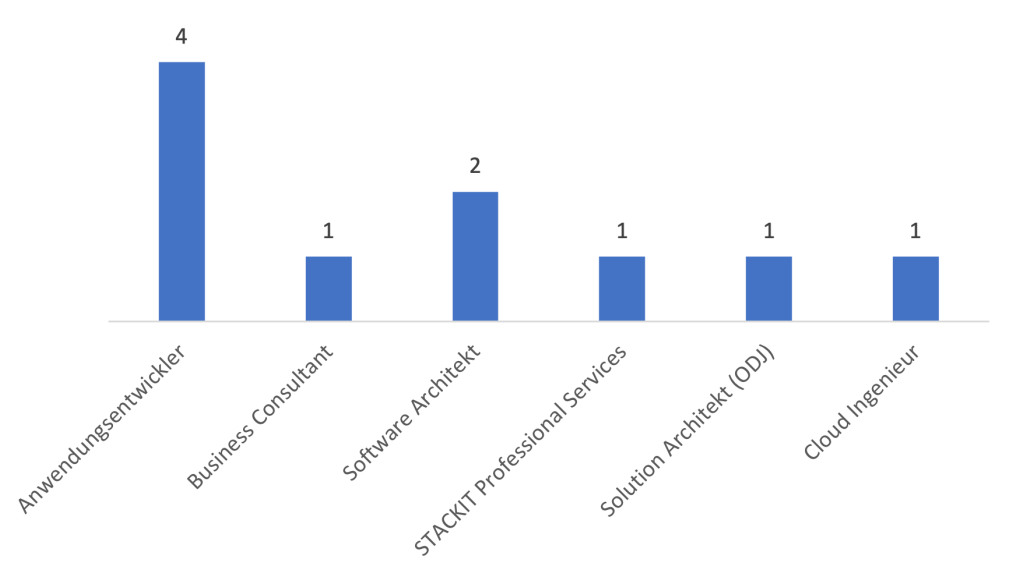

Abbildung 3.1: Verteilung der Arbeitsbereiche der Interviewteilnehmer

begonnen, dann werden die Fragen während dem Interview geschlossener und gegen Ende wieder offener. Welcher Ansatz sich am besten eignet, hängt sehr stark von den Fragen ab. Generell ist es wichtig in einer sinnvollen Reihenfolge durch das Interview zu führen und die Fragen nicht einfach zufällig zu sortieren. In Abschnitt 3.2.4 wird unter anderem die Reihenfolge der Fragen in diesem Interview vorgestellt. Es wurde der Timeglass-Ansatz gewählt. Dies liegt daran, dass sich dieser für die gewählten Fragen am besten geeignet hat und für eine sinnvolle Struktur sorgt.

3.2.3 Teilnehmer des Experteninterviews

Bei der Auswahl der Teilnehmer war das Ziel, möglichst viele Informationstechnologie (IT) Bereiche abzudecken, um möglichst viele verschiedene Ansichten mit einzubeziehen. Da diese Arbeit in Kooperation mit der SIT entsteht, wurden Mitarbeiter aus den verschiedensten Bereichen der SIT und ihrer Dienstleister herangezogen. Das Ziel war zunächst 10 Teilnehmer anzufragen und bei Bedarf noch mehr Interviews zu führen. Dies hing davon ab, ob nach 10 Teilnehmern weiterhin neue Erkenntisse gefunden werden. Da das nicht der Fall war, blieb es bei insgesamt 10 Teilnehmer. Acht von der SIT und zwei von Camao Tec, einem Software Dienstleister mit dem die SIT zusammenarbeitet.

Wie in Abbildung 3.1 zu sehen ist, wurden die 10 Teilnehmer aus verschiedensten Bereichen ausgewählt. Bei den Anwendungsentwicklern handelt es sich um sehr erfahrene Entwickler auf Professional bzw. Senior Professional Niveau. Zwei der vier Anwendungsentwickler sind dabei von Camao Tec. Generell wurde darauf geachtet, Mitarbeiter mit möglichst vielfältiger Erfahrung, auch bei anderen Unternehmen, auszuwählen. Bei der STACKIT handelt es sich um eine hauseigene, deutsche Cloudlösung der SIT. Der Mitarbeiter von STACKIT Professional Services unterstützt dabei Kunden ihre Lösungen in der Cloud zu betreiben.

3.2.4 Fragen des Experteninterviews

In der folgenden Aufzählung werden die Fragen des Interviews vorgestellt und diskutiert. Diese wurden mit Hilfe der in Abschnitt 3.2.1 vorgestellten Ziele des Experteninterviews erstellt. Die ersten Fragen beschäftigen sich damit die Arbeit zu motivieren und zu ermitteln welche Arten von sicherheitskritischen Daten in der Praxis existieren. Anschließend geht es vor allem darum Anforderungen für das Decision Support System zu sammeln. Dafür werden auch Tools herangezogen, die die Experten aktuell als Informationsquelle verwenden.

Frage 1: *Wenn Sie ein Dokument zum Thema „sicherheitskritische Daten in der Cloud“ lesen würden, welche Informationen würden Ihnen bei Ihrer Arbeit helfen?*

Bei dieser Frage geht es darum herauszufinden, welche Punkte besonders relevant für diese Arbeit sind. Der Interviewpartner wird die Stichpunkte angeben, die ihn besonders zu diesem Thema interessieren und hilfreich wären. Dadurch kann sichergestellt werden, dass die Ergebnisse dieser Arbeit einen Mehrwert bieten.

Frage 2: *Welche Arten von sicherheitskritischen Daten gibt es Ihrer Meinung nach? Wie lassen sich diese kategorisieren?*

Bei dieser Frage soll ermittelt werden, welche Arten von sicherheitskritischen Daten in der Praxis existieren und damit in dieser Arbeit behandelt werden müssen.

Frage 3: *In welchen Schritten eines Softwareprojektes spielen die Anforderungen für die Verarbeitung von sicherheitskritischen Daten eine Rolle? (Mehrfachauswahl möglich)*

- *Architektur*
- *Entwicklung*
- *Betrieb*

Bei dieser Frage soll die Zielgruppe dieser Arbeit ermittelt werden. Für wen das Thema sicherheitskritische Daten in der Cloud überhaupt interessant ist und an wen sich die ermittelten Lösungen richten.

Frage 4: *Was ist Ihre Hauptquelle um Informationen zur Verarbeitung von sicherheitskritischen Daten zu bekommen? (Mehrfachauswahl möglich)*

- *Firmeninterne Dokumente (Vorgaben)*
- *Wissenschaftliche Arbeiten*
- *Stack Overflow*
- *Blog-Beiträge*
- *White-Papers*
- *Dokumentation der eingesetzten Tools/Services*
- *Andere*

Bei dieser Frage geht es darum zu überprüfen, ob eine MLR in der Praxis Sinn macht. Denn sollten die Interviewpartner bei dieser Frage angeben, dass wissenschaftliche Arbeiten ihre Hauptquelle darstellt und andere Quellen so gut wie gar nicht verwendet werden, macht es keinen Sinn eine MLR gegenüber einer SLR durchzuführen. Dafür ist hier eine Vorauswahl gegeben, die vom Interviewpartner bei Bedarf auch um weitere Möglichkeiten ergänzt werden kann.

Frage 5: *Wie schwer ist nach Ihrem Empfinden die Suche nach Informationen zu diesem Thema?*

- *Sehr schwer*
- *Schwer*
- *Weder schwer noch einfach*
- *Einfach*
- *Sehr einfach*

Das Ziel dieser Frage ist zu ermitteln, wie hoch der Bedarf an einem System ist, das einen bei der Suche nach Informationen zum Thema sicherheitskritischen Daten in der Cloud unterstützt und somit hilft richtige Entscheidungen zu treffen.

Frage 6: *Welche Websites verwenden Sie um Informationen zu diesem Thema zu finden?*

- *Google*
- *Google Scholar*
- *Bing*
- *DuckDuckGo*
- *Andere*

Es soll geprüft werden, welche Webseiten die Interviewpartner bisher verwenden um Informationen zum Thema zu finden. Als Beispiele sind hier einige Suchmaschinen aufgelistet. Neunennungen sind hier ebenfalls möglich um herauszufinden, ob es eventuell schon eine Website gibt, die dieses Problem löst.

Frage 7: *Fühlen Sie sich von diesen Websites bei der Suche nach den Informationen ausreichend unterstützt?*
Wenn ja, welche Funktionen unterstützten Sie hier besonders?

Bei dieser Frage wird ermittelt, ob eine der genannten Webseiten den Nutzer bereits gut unterstützen, um Informationen zu finden und wenn das der Fall sein sollte, welche Funktionen hier besonders positiv herausstechen. Dadurch wird ermittelt, welche Funktionen Teil der umgesetzten Lösung sein sollten, die im Rahmen dieser Arbeit erstellt wird.

Frage 8: *Welche Probleme gibt es bei der Suche nach relevanten Informationen bei diesen Websites?*

Es wird geprüft, welche Probleme bei der Nutzung der bisherhigen Lösungen bestehen. Dadurch soll herausgefunden werden, welche Dinge bei der umgesetzen Lösung vermieden werden sollten bzw. welche Funktionen umgesetzt werden sollten, um diese Probleme zu vermeiden.

Frage 9: *Was könnten die Websites besser machen, damit Sie die Informationen die Sie benötigen einfacher finden?*

Hierbei handelt es sich um eine Frage, die sich an die Kreativität der Interviewpartner richtet. Es sollen hier zusätzliche Anforderungen gefunden werden, die das Decision Support System bereichern und die Zufriedenheit verbessern können.

Frage 10: *Ist der rechtliche Aspekt von sicherheitskritischen Daten relevant in Ihrer täglichen Arbeit? Warum?*

Bei dieser Frage geht es darum zu prüfen, wie wichtig es ist den rechtlichen Aspekt von sicherheitskritischen Daten in dieser Arbeit zu behandeln.

Frage 11: *Welche Herausforderungen sehen Sie beim Verarbeiten von sicherheitskritischen Daten in der Cloud?*

Es soll ermittelt werden, welche Themen bei der MLR behandelt werden sollten und wie die Forschungsfragen formuliert werden sollen. Dadurch wird sichergestellt, dass die MLR zielgerichtet ist und Lösungen für die gegebenen Probleme liefert.

Frage 12: *Wie unterscheidet sich, Ihrer Meinung nach, die Schwarz Gruppe von anderen Unternehmen was den Umgang mit sicherheitskritischen Daten in der Cloud betrifft?*

Es soll erhoben werden, wodurch sich die Anforderung an Sicherheit bei einem Konzern, wie der Schwarz Gruppe, von anderen Unternehmen unterscheidet. Aus diesem Grund wurden geziehlt Interviewpartner mit vielfältiger Erfahrung auch bei anderen Unternehmen ausgewählt.

Frage 13: *Zum Thema Cloud und Datenverarbeitung: Was sind für Sie noch offene Fragen die Sie gerne beantwortet hätten?*

Bei der letzten Frage geht es darum zu prüfen, ob es noch zusätzlich offne Fragen gibt, die noch in der Arbeit betrachtet werden sollten. Dadurch kann ebenfalls geprüft werden, ob alle relevanten Themen bereits bei den anderen Fragen behandelt wurden.

3.2.5 Ergebnisse des Experteninterviews

In der folgenden Auflistung wird für jede gestellte Frage das Ergebnis präsentiert und diskutiert.

Frage 1: *Wenn Sie ein Dokument zum Thema „sicherheitskritische Daten in der Cloud“ lesen würden, welche Informationen würden Ihnen bei Ihrer Arbeit helfen?*

In Abbildung 3.2 ist das Ergebnis der ersten Frage des Interviews zu sehen. Bei dieser Frage handelt es sich um eine offene Frage. Aus diesem Grund gab es eine große Anzahl an verschiedenen Antworten, bei denen es häufige Überschneidungen gab. *Best Practices* wurden beispielsweise mit acht Nennungen von fast jeder interviewten Person genannt. Generell fällt auf, dass mit Nennungen wie *Best Practices*, *Zu beachtende Dinge* und *Nennung von Standards* das Verlangen nach grundlegenden Kenntnissen sehr hoch ist. Hiermit kann die Notwendigkeit einer Lösung zur Erlangung und Verteilung von grundlegenden Kentnissen belegen, zum Beispiel auf Grundlage einer Pattern-Sprache. Durch die häufige Nennung von *Arten von sicherheitskritischen Daten* fällt zusätzlich

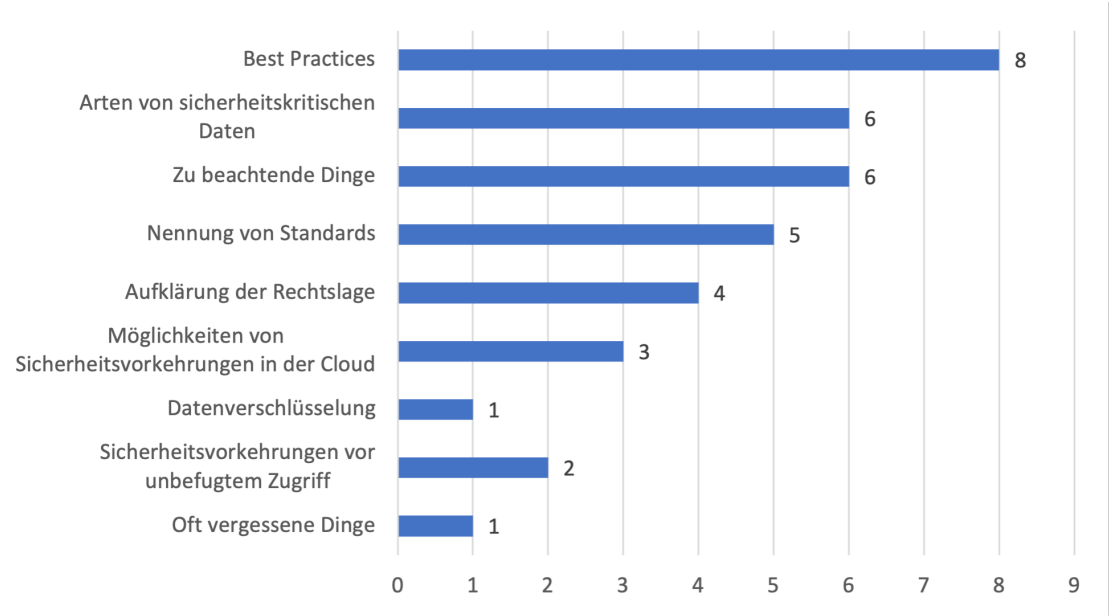

Abbildung 3.2: Ergebnis Frage 1 des Interviews

auf, dass Bedarf für eine Definition des Begriffs der *sicherheitskritischen Daten* besteht. Ein weiterer Punkt, der viermal genannt wurde, ist die *Aufklärung der Rechtslage*. Es scheint hier noch Bedarf an Aufklärung zu bestehen. Die weniger häufig genannten Punkte wie *Datenverschlüsselung* oder *Oft vergessene Dinge* lassen sich unter *Best Practices* subsumieren.

Frage 2: *Welche Arten von sicherheitskritischen Daten gibt es Ihrer Meinung nach? Wie lassen sich diese kategorisieren?*

Bei dieser Frage waren keine Antwortmöglichkeiten vorgegeben. Wie in Abbildung 3.3 zu sehen ist, wurden zwei Arten von sicherheitskritischen Daten von allen genannt: *personenbezogene/personenbeziehbare Daten* und *geschäftskritische Daten*. Die Nennung von personenbezogene bzw. personenbeziehbare Daten wurden von den Experten damit begründert, dass es sich dabei um rechtlich relevante Daten handelt. Die DSGVO, die innerhalb der EU gilt, wurde hier ebenfalls erwähnt.

Bei geschäftskritischen Daten handelt es sich um Daten, die bei Verlust dem Unternehmen finanziell schaden könnten und aus diesem Grund besonders geschützt werden sollten. Bei der Schwarz Gruppe ist das durch Lidl und Kaufland ein häufiges Thema. Daten, wie Einkaufspreise, können bei Veröffentlichung einen massiven Wettbewerbsnachteil nach sich ziehen.

Des Weiteren wurden von vier Experten *authentifizierungs und autorisierungs Daten (Nutzer/Server)* genannt. Hierbei wurden Einloggdaten von Nutzern für Anwendungen und Einloggdaten für Systeme, wie Datenbanken, zusammengefasst. Auch diese sind laut den Experten sehr schützenswert, da sie durch den öffentlichen Bereich der Cloud eine sehr wichtige Sicherheitsbarriere darstellen. Sollte jemand außerhalb des Unternehmens Zugriff auf die Einloggdaten eines Nutzers bekommen und damit auf eine Anwendung

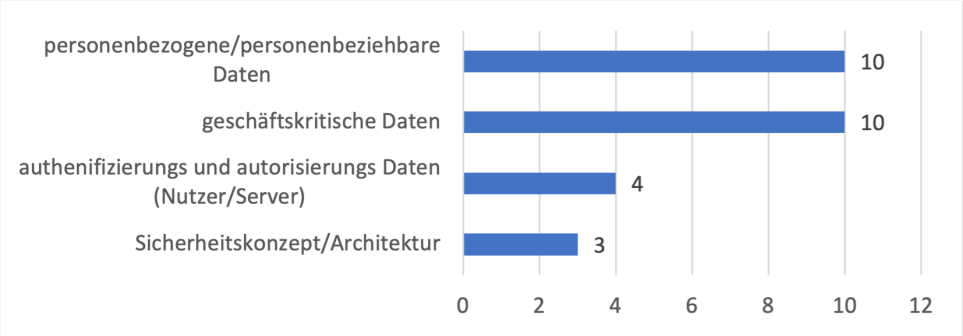

Abbildung 3.3: Ergebnis Frage 2 des Interviews

zugreifen, kann dieser eventuell sehr einfach an geheime Daten gelangen. Ebenso gilt das für technische Einloggdaten. Da bei der Cloud auf die Infrastruktur von überall zugegriffen werden kann, könnte man mit den entsprechenden Zugangsdaten direkt auf Datenbanken zugreifen und an geheime Daten gelangen.

Eine Art von sicherheitskritischen Daten die nur von drei der interviewten Personen genannt wurde, ist das *Sicherheitskonzept* bzw. die *Architektur*. Dies ist laut den Interviewpartnern zusätzlich eine schützenswerte Information, da bei allgemeiner Kenntnis sehr leicht bekannte Sicherheitslücken ausgenutzt werden können. Es wurde aber auch stark betont, dass die Geheimhaltung auf keinen Fall ein gutes Sicherheitskonzept ersetzt, sondern nur eine zusätzliche Hürde für einen Angreifer darstellt. Aus diesem Grund werden in dieser Arbeit die drei meist genannten Daten behandelt: *personenbezogene/personenbeziehbare Daten*, *geschäftskritische Daten* und *authentifiszierungs und autorisierungs Daten (Nutzer/Server)*.

Frage 3: *In welchen Schritten eines Softwareprojektes spielen die Anforderungen für die Verarbeitung von sicherheitskritischen Daten eine Rolle? (Mehrfachauswahl möglich)*

Bei dieser Frage gab es drei vorgegebene Antworten: *Architektur*, *Entwicklung* und *Betrieb*. Hier wurden alle Antworten von allen Interviewteilnehmern angegeben. Als Grund dafür wurde von allen genannt, dass wenn einer der Schritte keine Vorkehrungen für sicherheitskritische Daten treffen würde, immer das gesamte System anfällig wäre, egal wie gut die Arbeit der anderen Schritte war.

Frage 4: *Was ist Ihre Hauptquelle um Informationen zur Verarbeitung von sicherheitskritischen Daten zu bekommen? (Mehrfachauswahl möglich)*

Bei dieser Frage waren folgende Antworten vorgegeben: *Firmeninterne Dokumente (Vorgaben)*, *Wissenschaftliche Arbeiten*, *Stack Overflow*, *Blog-Beiträge*, *White-Papers* und *Dokumentation der eingesetzten Tools/Services*. Es war aber auch möglich, weitere Quellen anzugeben.

Von den angegebenen Antwortmöglichkeiten, wurde *Stack Overflow* von keiner der interviewten Personen genannt. In Abbildung 3.4 ist zu sehen, dass die dominante Hauptquelle aller Befragten interne Dokumente bzw. Vorgaben sind. Die zweite meist genannte Quelle sind Dokumentationen der eingesetzten Tools/Services. Hier wurde vor allem auf die Dokumentationen der großen Cloudanbieter, wie Amazon Web Services (AWS), Google

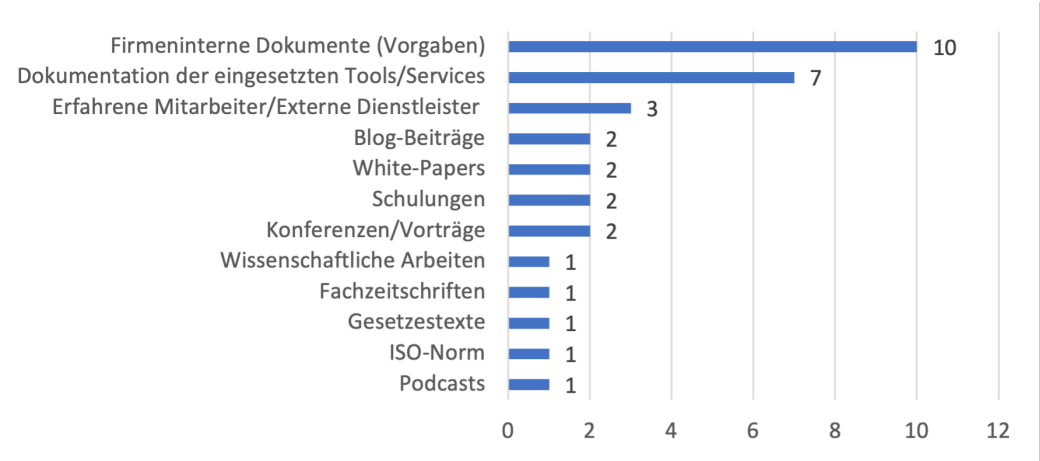

Abbildung 3.4: Ergebnis Frage 4 des Interviews

Cloud (GC) und Microsoft Azure (MSA) Bezug genommen. Des Weiteren gab es noch einige weitere seltene Nennungen, wie *Podcasts* und *ISO-Norm*. Wissenschaftliche Arbeiten wurden nur von zwei Personen als Hauptquelle genannt. Bei Nachfrage bei den anderen Personen wurde die schwierige Zugänglichkeit von wissenschaftlichen Arbeiten als Grund für die Nichtnennung angegeben.

Als Grund für die hohe Anzahl an verschiedenen Quellen wurde die Vielseitigkeit des Themas genannt. Meist enthält eine Quelle nicht alle Informationen die notwendig sind. Hier wurde beispielsweise davon berichtet, dass interne Dokumente vorgeben, welche Sicherheitsvorkehrungen getroffen werden müssen, allerdings dann konkrete Umsetzungen meist in den Dokumentationen der eingesetzen Tools zu finden sind. Konferenzen, Schulungen und Podcasts wurden vor allem als allgemeine Wissensquellen genannt, die zum Aufbau von Kenntnissen zum Thema verwendet werden. Es lässt sich an den Antworten der Experten erkennen, dass Bedarf an einer zentralen Lösung besteht, die alle genannten Quellen integriert.

Frage 5: *Wie schwer ist nach Ihrem Empfinden die Suche nach Informationen zu diesem Thema?*

Bei dieser Frage war eine Likert-Skala gegeben, um zu bestimmen, wie schwierig es ist, Informationen zum Thema sicherheitskritische Daten in der Cloud zu finden. Die durchschnittliche Antwort lag, wie in Abbildung 3.5 zu sehen ist, leicht über schwer. Es haben insgesamt acht der Teilnehmer schwer angegeben und zwei sehr schwer. Dadurch ist zu sehen, dass es aktuell noch schwer fällt, Informationen mit den vorhandenen Systemen zu finden. Es besteht also Bedarf an dem im Rahmen dieser Arbeit entstehenden Decision Support System.

Frage 6: *Was ist Ihre Hauptquelle um Informationen zur Verarbeitung von sicherheitskritischen Daten zu bekommen? (Mehrfachauswahl möglich)*

Bei dieser Frage waren folgende Suchmaschinen als Antwortmöglichkeiten vorgegeben: *Google*, *Google Scholar*, *Bing*, *DuckDuckGo*. Es war zudem möglich, weitere Websites anzugeben. Wie in Abbildung 3.6 zu sehen ist, wurde mit großer Mehrheit Google als Antwort angegeben. Der Grund hierfür war, dass für viele intuitiv Google die erste

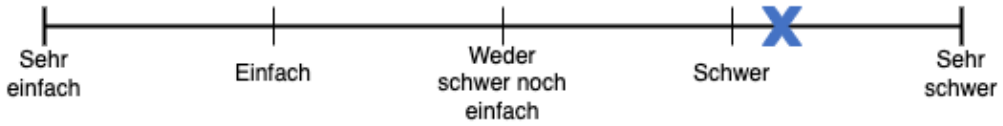

Abbildung 3.5: Ergebnis Frage 5 des Interviews

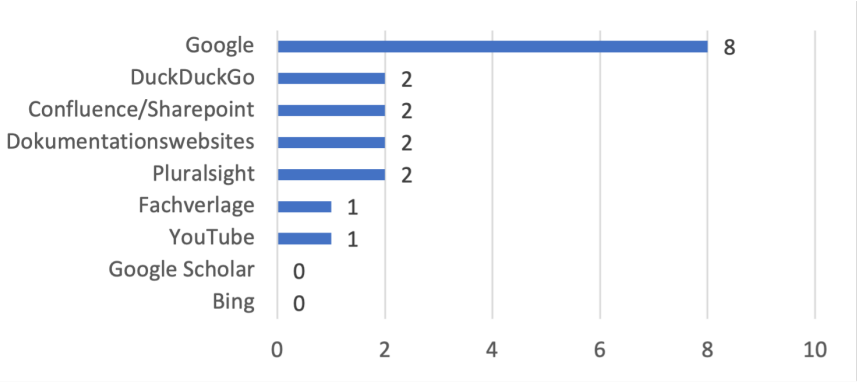

Abbildung 3.6: Ergebnis Frage 6 des Interviews

Anlaufstelle ist, um Informationen zu finden. Von zwei Personen wurde aus Gründen von persönlicher Präferenz DuckDuckGo angegeben. Diese kommt dadurch, dass dort keine personalisierte Suche durchgeführt wird. Ebenfalls von zwei Nutzern wurden interne Dokumentationswerkzeuge (Confluence, etc.) als Quelle angegeben. Confluence und Sharepoint werden intern bei der Schwarz Gruppe zur Dokumentation verwendet. Zusätzlich wurden auch Dokumentationswebsites von Anbietern wie AWS, GC und MSA angegeben.

Generell fällt auf, dass keine Website angegeben wurde, die sich speziell mit sicherheitskritischen Daten in der Cloud auseinandersetzt und Lösungen für häufig auftretende Probleme anbietet. Das bedeutet, dass bisher kein Produkt existiert wie das, in dieser Arbeit umgesetzt wird.

Frage 7: *Fühlen Sie sich von diesen Websites bei der Suche nach den Informationen ausreichend unterstützt?*
Wenn ja, welche Funktionen unterstützten Sie hier besonders?

Bei dieser Frage waren sich die interviewten Personen trotz des, in Abbildung 3.7 zu sehenden, Ergebnis sehr einig. Sechs der zehn Teilnehmer gaben nein an und vier ja, diese allerdings betonten, dass sie die Leistung tatsächlich nur als ausreichend bezeichnen würden, was einer vier als Schulnote entsprechen würde. Das bedeutet, dass keine der in Frage 6 genannten Websites die Teilnehmer gut bei der Suche untersützt. Aus diesem Grund wurden auch nicht besonders viele Funktionen genannt, die den Nutzern helfen.

Bei Google wurde positiv hervorgehoben, dass es dort möglich ist, das Alter von Ergebnissen zu begrenzen, beispielsweise auf ein Jahr. Dadurch werden die Ergebnisse gerade bei einem Thema, das sich so schnell entwickelt wie Sicherheit, stark auf die aktuellsten

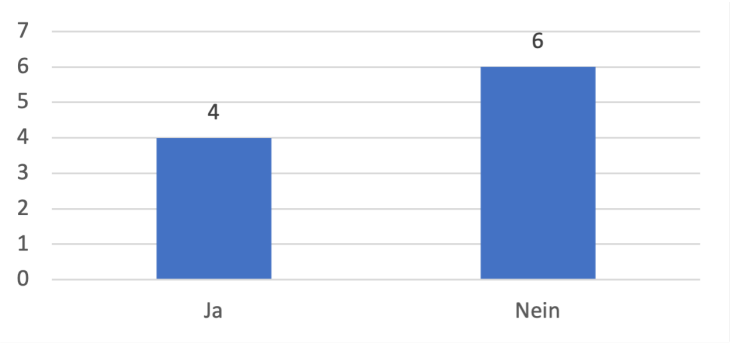

Abbildung 3.7: Ergebnis Frage 7 des Interviews

eingeschränkt. Bei DuckDuckGo wurde angegeben, dass die nicht personalisierte Suche bei der Präzision der Ergebnisse einer Suchanfrage deutlich verbessern würde. Zudem ist es dort auch möglich die Sprache einzustellen in denen die Ergebnisse dokumentiert sein sollen. Was in einem englisch dominierten Bereich auch von Vorteil ist. Allgemein wurde auch darauf hingewiesen, dass Querverweise, also Verbindungen zu weiteren passenden Themen, generell sehr hilfreich sind.

Frage 8: *Welche Probleme gibt es bei der Suche nach relevanten Informationen bei diesen Websites?*

Bei dieser Frage gab es durch die Unzufriedenheit bei Frage 7 sehr viele Antworten, die sich allerdings sehr stark auf einige wenige Funktionen fokussiert haben. Da sich die Antworten auf bestimmte Webseiten bezieht, sind diese in Abbildung 3.8 entsprechend markiert. Dabei steht (G) für Google und (C/S) für Confluence/Sharepoint. Da nur für diese Websites Probleme angegeben wurden, wird sich hier auch auf diese beschränkt. Wie in Abbildung 3.8 zu sehen ist, stellen vor allem die Suchbegriffe ein großes Problem dar. Hier wurde davon berichtet, dass es oft sehr schwierig ist, die richtigen Suchbegriffe zu finden, um zielgerichtete Ergebnisse zu bekommen. Dies liegt aber wohl auch in der Natur von einer allgemeinen Suchmaschine. Generell ist zu sehen, dass bei Google sehr oft die Suche und die Ergebnisse das Problem sind. Es wurde genannt, dass diese zu viele oder auch veraltet sein können. Es ist auch schwierig, deren Qualität richtig einzuschätzen. Denn gerade bei Blog-Posts weiß man oft nicht, wie vertrauenswürdig die Quelle ist. Bei Confluence bzw. Sharepoint gibt es vor allem organisatorische Probleme, wie Zugriffsrechte oder teilweise schlechte Suche durch Tags. Generell besteht auch das Problem, dass jedes Team bzw. Projekt ihre eigene Seite pflegt und diese nicht zentral organisiert sind. Dadurch kommt es häufig vor, dass die gleichen Themen von verschiedenen Teams erarbeitet werden.

Grundlegend liegen vor allem die Probleme darin, dass es keine zentrale und spezialisierte Lösung gibt, die wichtige Informationen über sicherheitskritische Daten in der Cloud einfach und übersichtlich dokumentiert.

Frage 9: *Was könnten die Websites besser machen, damit Sie die Informationen die Sie benötigen einfacher finden?*

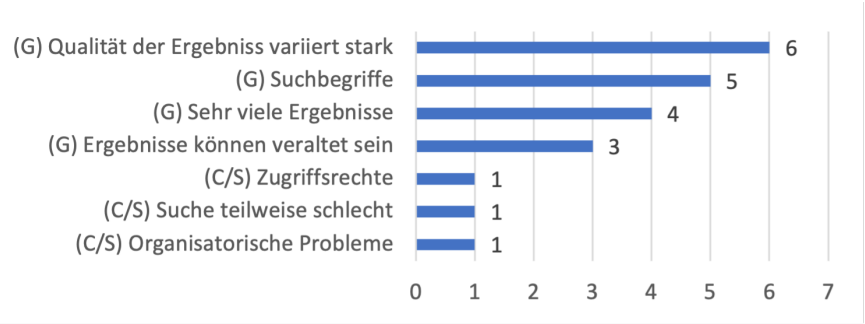

Abbildung 3.8: Ergebnis Frage 8 des Interviews

Bei dieser Frage ging es darum, Funktionen zu finden, die sich Nutzer beim Decision Support System wünschen würden. Dabei wurden folgende Funktionen genannt (Zahlen in Klammern kennzeichnen Mehrfachnennungen mit der entsprechenden Anzahl und das Präfix in Klammern kennzeichnet auf welche Website die Funktion bezogen ist):

- (Allgemein) Grundlegende Kategorisierung (5)
- (Allgemein) Best Practices / Patterns an denen man sich entlang hangeln kann (3)
- (Allgemein) Tags/Dashboards für ersetzte/veraltete Lösungen (3)
- (Allgemein) Generelle Übersicht/Einstieg über das Thema (3)
- (Allgemein) Filter nach Technologie (2)
- (Pluralsight) Benötigtes Vorwissen angeben um Artikel/Video zu verstehen
- (Blogs) Vollständige und sinnvolle Inhaltsangaben
- (Confluence/Sharepoint) Intelligente Suche
- (Confluence/Sharepoint) Eine zentrale Lösung

Generell sieht man, dass gut strukturierte Best Practices und Patterns gewünscht werden, die sinnvoll miteinander verknüpft werden. Dies entspricht einer Pattern-Sprache, die die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen, behandelt. Es ist zu sehen, dass vor allem eine gute Filterung und Tags notwendig sein werden. Zudem sollte der Einstieg möglichst einfach gestaltet werden, da dies bei einem neuen Thema besonders schwierig ist, wenn man direkt mit einer großen Menge an Daten konfrontiert wird.

Frage 10: *Ist der rechtliche Aspekt von sicherheitskritischen Daten relevant in Ihrer täglichen Arbeit? Warum?*

Wie in Abbildung 3.9 zu sehen ist, ist für 80 Prozent der rechtliche Aspekt ein Teil ihrer täglichen Arbeit. Alle von diesen Befragten begründeten das damit, dass jeder Mitarbeiter in der Verantwortung steht, rechtliche Rahmenbedingungen selbst einzuhalten und darauf hinzuweisen, sollten diese nicht eingehalten werden. Es wurde auch erwähnt, dass bei Nichteinhaltung der Datenschutzgesetze erhebliche Strafen drohen. Zwei der Befragten beantworteten die Frage mit Nein. Die Nein-Antworten lassen sich nicht exklusiv auf

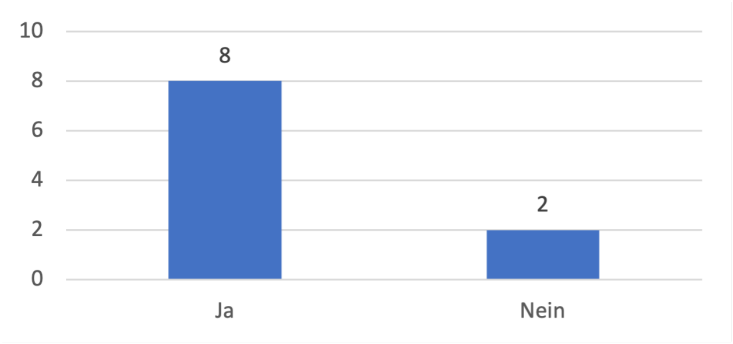

Abbildung 3.9: Ergebnis Frage 10 des Interviews

einen Bereich aus dem die Experten kommen zurückführen. Sie wurden damit begründet, dass es sich bei rechtlichen Themen um fachliche Anforderungen handelt und aus diesem Grund nicht Teil ihrer täglichen Arbeit sind. Sie werden nur dann Teil ihrer Arbeit, wenn sie angefordert werden.

Es lässt sich daraus folgern, dass der rechtliche Aspekt eine große Rolle spielt und damit auch in dieser Arbeit betrachtet werden muss.

Frage 11: *Welche Herausforderungen sehen Sie beim Verarbeiten von sicherheitskritischen Daten in der Cloud?*

Bei dieser Frage wurden folgende Stichpunkte genannt (Zahlen in Klammern repräsentieren Mehrfachnennungen):

- Zu wenig Wissen über das Thema vorhanden (4)
- Eine vollkommene Sicherheit gibt es nicht (sollte jedem klar sein) (2)
- Komplexität in der Cloud (2)
- Zu wenig Informationen einfach/schnell verfügbar
- Welche Cloud/Cloud-Angebot erfüllen die Vorgaben?
- Schutz der eigenen Systeme und deren Kommunikation
- Aktualität der Lösungen
- Man darf keine Fehler machen
- Einhaltung der DSGVO

Hier sind vor allem zwei Problematiken aufgelistet. Eines davon ist Kommunikation und Verständnis über das Thema innerhalb eines Unternehmens. Oft fehlt es an Kenntnissen oder Verständnis für das Thema Sicherheit. Das heißt, hier muss über das Thema informiert werden und ein einfacher, verständlicher Einstieg in das Thema angeboten werden. Die zweite Problematik sind Informationen. Sie sind meist nicht einfach zu finden und dann meist nicht vollständig. Hier wurde auch erwähnt, dass eine Verknüpfung zwischen

Vorgaben und Cloud Angebote komplett fehlt bzw. dieses Wissen nicht einfach abrufbar ist. Das schließt auf eine Lösung, die alle Fragen zum Thema sicherheitskritische Daten in der Cloud beantwortet und auch zentral regelt.

Frage 12: *Wie unterscheidet sich, Ihrer Meinung nach, die Schwarz Gruppe von anderen Unternehmen was den Umgang mit sicherheitskritischen Daten in der Cloud betrifft?*

Bei dieser Frage wurden folgende Stichpunkte genannt (Zahlen in Klammern repräsentiert Mehrfachnennungen):

- Unterscheidet sich nicht stark zu anderen Unternehmen in Deutschland
  - Starke Datenschutzbedenken
  - Sicherheitsbedenken der Daten (4)
- Unterschiedliche Sichtweisen zu Cloud in verschiedenen Bereichen
- Alles sehr stark geregelt (4)
- Offenheit für Cloud ist aber da (4)
- Eigene Cloud (STACKIT) (4)
- Zu wenig Erfahrung/Wissen im Unternehmen

Generell ist also zu sehen, dass es sich bei der Schwarz Gruppe gerade bei Bedenken der Cloud um einen klassischen europäischen Konzern handelt. Denn gerade der Betrieb von geschäftskritischen Daten wird sehr ernst genommen und stark geregelt. Zudem gelten auch sehr strenge Datenschutzgesetze, die ebenfalls eingehalten werden müssen. Grundlegend bleibt die Schwarz Gruppe sehr Cloud-offen, da sie mit der STACKIT eine eigene Cloud betreibt. Grundlegend lässt sich also folgern, dass es hohe Datenschutzanforderungen gibt und somit das Ergebnis auch in jeder Konzernstruktur einsetzbar ist.

Frage 13: *Zum Thema Cloud und Datenverarbeitung: Was sind für Sie noch offene Fragen die Sie gerne beantwortet hätten?*

Bei dieser Frage handelt es sich um eine offene Abschlussfrage, die vor allem darauf abzielt, nicht gestellte Fragen zu ermitteln. Keiner der Befragten hatte an dieser Stelle noch offene Punkte, die es zu besprechen gab. Dadurch kann gefolgert werden, dass die vorher gestellten Fragen das Thema sehr gut abgedeckt haben.

3.2.6 Formulierung der Forschungsfragen

Im Folgenden werden basierend auf den in Abschnitt 3.2.5 dokumentierten Experteninterviews die Forschungsfragen formuliert, die mit der MLR beantwortet werden sollen. Folgende Forschungsfragen ergaben die Interviews:

Forschungsfrage 1: *Welche Lösungen zum Schutz unterschiedlicher sicherheitskritischer Daten gibt es und wie kann man diese klassifizieren?*

Diese Frage ergibt sich vor allem aus Frage 1 (*Wenn Sie ein Dokument zum Thema „sicherheitskritische Daten in der Cloud“ lesen würden, welche Informationen würden Ihnen bei Ihrer Arbeit helfen?*) des Interviews. Hier wurde klar, dass eine der wichtigsten Punkte konkrete Lösungen für die verschiedenen Arten von sicherheitskritischen Daten sind. Der Teil mit der Klassifizierung ergibt sich vor allem aus Frage 9 (*Was könnten die Websites besser machen, damit Sie die Informationen die Sie benötigen einfacher finden?*). Dort ging es stark um Tags und geeignete Filter. Aus diesem Grund ist es wichtig, die Lösungen entsprechen zu klassifizieren und mit passenden Tags zu versehen.

Forschungsfrage 2: *Welche rechtlichen Anforderungen gibt es bei der Verarbeitung von sicherheitskritischen Daten in der Cloud?*

Diese Forschungsfrage wird aufgrund der Antworten zu Frage 10 (*Ist der rechtliche Aspekt von sicherheitskritischen Daten relevant in Ihrer täglichen Arbeit? Warum?*) gestellt. Dort wurde klar, dass die rechtlichen Anforderungen mit behandelt werden müssen, um mehr Kenntnisse in diesem Bereich zu sammeln.

Forschungsfrage 3: *Wie kann ein Unternehmen mithilfe eines Decision Support Systems bei der Suche nach geeigneten Lösungen unterstützt werden?*

Diese Frage ergibt sich einerseits aus dem Ziel dieser Arbeit, ein Decision Support System umzusetzen, und andererseits aus Frage 12 (*Wie unterscheidet sich, Ihrer Meinung nach, die Schwarz Gruppe von anderen Unternehmen was den Umgang mit sicherheitskritischen Daten in der Cloud betrifft?*). Dort wurde klar, dass es ein sehr wichtiges Thema ist, diese Informationen innerhalb eines Konzerns zu verbreiten und eine einheitliche Strategie zu schaffen.

3.3 Suchprozess der MLR

Das Ziel der MLR liegt darin, ein Bild über den aktuellen Stand und eine Übersicht über die Art der Daten zu erhalten, um die Anforderungen für das Decision Support System möglichst praxisnah gestalten zu können. Für den Suchprozess wurde eine von Garousi [GFM19] abgeleitete Methode verwendet. Die Suche startet mit drei initialen Such-Strings. Es werden zwei verschiedene Suchmaschinen für die Suche verwendet. Google und Google Scholar. Bei beiden wird sich auf die Ergebnisse der ersten Seite bis maximal 10 Einträge beschränkt. Bei Google ist dabei zu beachten, dass keine Anzeigen als Ergebnis gewertet wurden, um nur unbezahlte Treffer auszuwerten. Nach den drei initialen Such-Strings wurden diese iterativ anhand der erzielten Ergebnisse ausgeweitet. Dies wurde so lange durchgeführt, bis die Forschungsfragen anhand der Ergebnisse beantwortet werden konnten.

Für die MLR wurden ingesamt 10 Such-Strings verwendet, die sich im oben beschriebenen iterativen Verfahren in insgesamt vier Durchgängen ergeben haben. Die Such-String waren die folgenden:

- protect user data
- protect company data
- protect authentication data
- web security best practices
- cloud security best practices
- cloud data security patterns
- API Key security patterns cloud
- dsgvo cloud data security
- cloud provider dsgvo requirements
- dsgvo cloud

Nach der Sammlung der Ergebnisse dieser Such-Strings wurden wenig bis kaum neue Erkenntnisse gefunden und aus diesem Grund keine weitere Suche durchgeführt.

3.4 Datenerhebung der MLR

Für die Datenerhebung wurde sich für eine vereinfachte Version der von Garousi [GFM19] vorgestellten Methode verwendet. Dies lässt sich damit begründen, dass das Ziel der MLR ist einen generellen Überblick über das Thema Datensicherheit in der Cloud zu erhalten.

Die Ergebnisse des in Abschnitt 3.3 vorgestellen Suchverfahrens wurden in einer Exceltabelle festgehalten. Hier wurden folgende Informationen erfasst:

- Überschrift

  Überschrift der Website oder wissenschaftlichen Arbeit.

- Link

  Direkter Link zur Ressource

- Such-String

  Such-String mit Hilfe dessen diese Ressource gefunden wurde.

- Voting

  Hier wurde festgehalten, wie relevant die Resource für die Beantwortung der Forschungsfrage ist. Dabei wurden Farben verwendet. Rot steht für irrelevant, Gelb für eventuell relevant, Grün für relevant und Violett für besonders relevant, dies wurde bei sehr vielversprechenden Resourcen vergeben.

- Tags

  Hier wurden besonders relevante Schlagwörter festgehalten, die sich aus den Resourcen ableiten lassen, wodurch einfacher Gemeinsamkeiten zwischen den Resourcen feststellbar sind.

Dies wurde für Google und Google Scholar Ergebnisse in getrennten Tabellen festgehalten.

3.5 Datensynthese der MLR

Um die Daten aus den verschiedenen Quellen zu erfassen, wurden zunächst das Voting als Richtwert herangezogen und alle Quellen die ein gelbes oder rotes Voting erhalten haben aussortiert. Anschließend wurden die verschiedenen Quellen in die passende Forschungsfrage gruppiert. Dies wurde mit Hilfe der dokumentierten Schlagwörter durchgeführt. Anschließend wurden die Quellen nach wiederauftretenden Informationen überprüft. In Abschnitt 3.6 wurden die Informationen festgehalten, die von den meisten Quellen aufgelistet wurden.

3.6 Ergebnisse der MLR

In den folgenden Abschnitten werden die Ergebnisse der MLR präsentiert. Dabei werden Ausschnitte mit kurzen Beschreibungen gegeben um Redundanzen mit Kapitel 4 zu reduzieren, denn in diesem Kapitel werden Pattern-Sprachen präsentiert, die ebenfalls Lösungen für die Forschungsfragen liefern.

3.6.1 Ergebnisse für die Forschungsfrage 1

Für die Forschungsfrage 1, vorgestellt in Abschnitt 3.2.6, hat sich bei der Review ergeben, dass Lösungen zum Schutz sicherheitskritischer Daten sich in zwei Klassen aufteilen lassen: *Lösungen für Nutzerdaten zur Authentifizierung und Autorisierung* und *Lösungen für geschäftskritische und personenbezogene Daten*. Im Folgenden wurden jeweils die vier meist genannten Lösungen kurz aufgelistet und beschrieben. Dafür wurde sich entschieden, da im Rahmen der Datensysthese aufgefallen ist, dass sehr viele dargestellte Lösungen Überschneidungen mit der in Abschnitt 4.3 beschriebenen *Data Security Pattern* Pattern-Sprache haben. Dadurch würden die Informationen in dieser Arbeit redundant aufgelistet werden.

Lösungen für Nutzerdaten zur Authentifizierung und Autorisierung

- Passwörter verschlüsseln

  Passwörter direkt ohne Verschlüsselung in einer Datenbank zu speichern, sorgt für ein großes Risiko, da sie bei einem unauthorisierten Zugriff lesbar sind. Aus diesem Grund sollten Passwörter immer mit einem sicheren Hash-Algorithmus unkenntlich gemacht werden.

  Quellen: [Rus17], [22e], [Job19], [Rah14], [22c], [Omo22], [17], [Lam21], [Nid22], [Jon22], [Che22], [22b], [Kri21], [19a]

- Aktivitätenüberwachung

  Wenn es dazu kommt, dass Angreifer unerlaubt Zugriff auf Anwendungen oder Daten erhalten, sollte dies auch erkannt werden. Aus diesem Grund empfielt es sich Mechanismen zu implementieren, die verdächtige Aktivitäten erkennen und sperren.

  Quellen: [Rus17], [sir], [Job19], [Rah14], [Coo22], [22c], [Nid22], [Jon22], [Che22], [22b], [Kri21], [22d], [Har22], [Har22]

- Zentrale Benutzerverwaltung

  Durch eine individuelle Benutzerverwaltung in jeder Anwendung kann es dazu kommen, dass Nutzer auf Daten Zugriff erhalten, auf die sie nicht zugreifen können sollten. Aus diesem Grund ist es sinnvoll ein zentrales Benutzerverwaltungssystem innerhalb eines Unternehmens umzusetzen. Dieses System stellt Anwendungen die Benutzerrollen eines Nutzers zentral zur Verfügung. Dadurch können Rollen deutlich übersichtlicher und sicherer verwaltet werden.

  Quellen: [Rus17], [sir], [Job19], [22c], [Omo22], [Nid22], [Jon22]

- Sichere Authentifizierungs-Protokolle

  Um zu vermeiden, dass es bei der Übertragung von Benutzeranmeldeinformationen zu Sicherheitslücken kommt sollten ausschließlich sichere Protokolle, wie OAuth 2.0, verwendet werden.

  Quellen: [Rus17], [sir], [22e], [Job19], [Rah14], [22c], [Bit21], [Lam21], [Nid22], [Jon22], [Har22]

Lösungen für geschäftskritische und personenbezogene Daten

- Zugriffsverwaltung

  Es sollte stark darauf geachtet werden, welche Nutzergruppen Zugriff auf welche Daten haben. Geschäftsdaten sollten nur von jenen Personen einsehbar sein, die darauf auch Zugriff haben dürfen.

  Quellen: [Rus17], [sir], [22e], [22c], [Omo22], [Jon22], [22b], [22d]

- Security Audits

  Einführung von zusätzlichen Sicherheitsprüfungen um leicht vermeidbare Sicherheitslücken zu vermeiden. Beispielsweise die Einführung von Snyk, das regelmäßig Abhängigkeiten auf bekannte Sicherheitslücken und Updates prüft.

  Quellen: [Job19], [22c], [Lam21], [Nid22], [Jon22], [Che22], [22b], [Kri21], [19a], [22d], [Har22]

- Speicherung

  Verschlüsselte Speicherung von geschäftskritischen Daten. Dadurch wird verhindert, dass geschäftskritische Daten bei unerlaubtem Datenbankzugriff lesbar sind.

  Quellen: [Rus17], [sir], [22e], [Job19], [Rah14], [Coo22], [22c], [Omo22], [17], [Lam21], [Jon22], [Che22], [22b], [Kri21], [19a], [Har22]

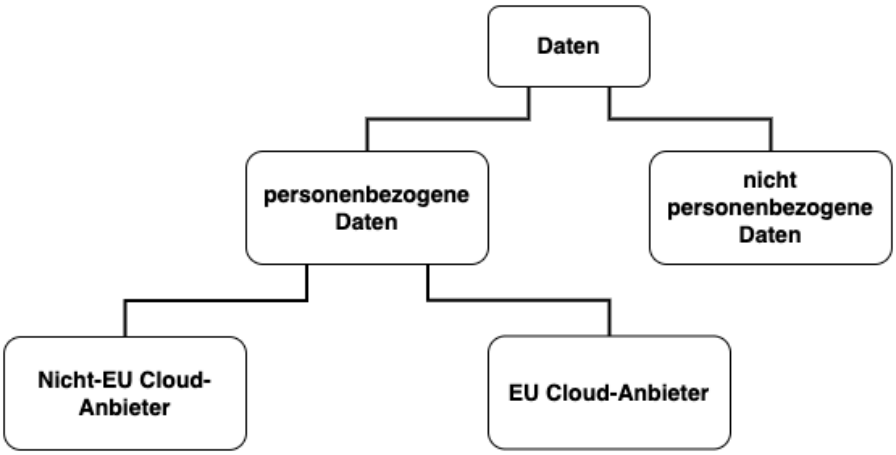

Abbildung 3.10: Visuelle Darstellung der verschiedenen rechtlichen Kategorien

- Daten Backups

  Um einen Totalverlust von Daten zu vermeiden, sollten in regelmäßigen Abständen Backups der Daten angelegt und sicher abgelegt werden.

  Quellen: [Rus17], [sir], [22e], [Job19], [21], [22c], [17], [Lim20]

3.6.2 Ergebnisse für die Forschungsfrage 2

Bei dieser Forschungsfrage, vorgestellt in Abschnitt 3.2.6, ging es darum zu prüfen, zu welchen technischen Lösungen man innerhalb der EU beim Umgang mit sicherheitskritischen Daten rechtlich verpflichtet ist. Bei sicherheitskritischen Daten werden in dieser Arbeit *geschäftskritische Daten*, *personenbezogene Daten* und *Nutzerdaten zur Authentifizierung und Autorisierung* untersucht.

Wie in Abbildung 3.10 zu sehen ist, lassen sich rechtlich Daten vereinfacht in personenbezogene und nicht personenbezogene Daten unterscheiden. Aus diesem Grund werden im folgenden die Daten in diese beiden Kategorien behandelt. *Nutzerdaten zur Authentifizierung und Autorisierung* werden hierbei in personenbezogene Daten zusammengefasst, da sie ebenfalls zu dieser Kategorie gehören. *Geschäftskritische Daten* gelten als nicht personenbezogene Daten.

Nicht personenbezogene Daten

Für nicht personenbezogene bzw. geschäftskritische Daten schreibt der Gesetzgeber keine technischen Anforderungen vor. Um geschäftskritische Daten allerdings unter rechtlichen Schutz zu stellen, sodass Personen bei Offenlegung erlangter Daten rechtlich belangt werden können, bedarf es nach dem Gesetz zum Schutz von Geschäftsgeheimnissen (GeschGehG) einem ausreichenden Schutz der Daten [19b]. Aus diesem Grund empfiehlt es sich diese gut zu schützen.

Personenbezogene Daten und Nutzerdaten zur Authentifizierung und Autorisierung

Die europäische Komission definiert personenbezogene Daten wie folgt:

„Personenbezogene Daten sind alle Informationen, die sich auf eine identifizierte oder identifizierbare lebende Person beziehen. Verschiedene Teilinformationen, die gemeinsam zur Identifizierung einer bestimmten Person führen können, stellen ebenfalls personenbezogene Daten dar. Personenbezogene Daten, die anonymisiert, verschlüsselt oder pseudonymisiert wurden, aber zur erneuten Identifizierung einer Person genutzt werden können, bleiben personenbezogene Daten und fallen in den Anwendungsbereich der Datenschutz-Grundverordnung. Personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann, gelten nicht mehr als personenbezogene Daten. Damit die Daten wirklich anonymisiert sind, muss die Anonymisierung unumkehrbar sein.“ [19d]

Diese Daten fallen unter die DSGVO und bedürfen besonderen Schutz, was Auswirkungen auf die Verarbeitung derer in der Cloud nach sich zieht. Denn die MLR ergab, wie in Abbildung 3.10 zu sehen ist, dass zwischen EU und nicht EU Cloud-Anbieter unterschieden werden muss.

Bei EU-Anbietern müssen die Daten, wie in der DSGVO vorgeschrieben, ausreichend geschützt werden. Es sollte also ein DSGVO zertifizierter Cloud-Anbieter verwendet werden und die Daten die dort verarbeitet werden, müssen ausreichend vor unerlaubtem Zugriff geschützt werden. Das bedeutet, dass dort die in Abschnitt 3.6.1 vorgestellten Maßnahmen getroffen werden müssen.

Bei nicht EU-Anbietern müssen zusätzliche Maßnahmen getroffen werden. Denn wie in der Ausarbeitung der wissenschaftlichen Dienste [Die21] festgestellt wurde, dürfen Daten nur an Drittländer ausgeliefert werden, wenn in diesem Land mindestens die gleichen Datenschutzstandards wie in der EU gelten. Ländern bei denen dies beispielsweise laut der Ausarbeitung nicht der Fall ist, sind die USA. Die besondere Nennung lässt sich auf die großen Anbieter, wie AWS, die aus den USA stammen zurückführen. Für diese Fälle müssen neben Zertifizierungen und Garantien der Anbieter zusätzliche Maßnahmen getroffen werden, die jeglichen Zugriff ausländischer Behörden auf personenbezogene Daten von EU-Bürgern zu jedem Zeitpunkt verhindern.

3.6.3 Ergebnisse für die Forschungsfrage 3

Zur Beantwortung dieser Forschungsfrage, vorgestellt in Abschnitt 3.2.6, wurden die Antworten aus dem in Abschnitt 3.2 vorgestellten Interviews verwendet. Speziell Frage 9 listet einige Anforderungen an das Decision Support System auf, darunter folgende Aussagen der Interviewpartner:

- Grundlegende Kategorisierung
- Best Practices / Patterns an denen man sich entlang hangeln kann
- Generelle Übersicht/Einstieg über das Thema
- Vollständige und Sinnvolle Inhaltsangaben
- Eine zentrale Lösung

Hier wird klar der Wunsch nach einer zentralen Lösung mit einer übersichtlichen Darstellung von Patterns und Pattern-Sprachen zu diesem Thema geäußert. Viele dieser Funktionen bietet bereits der *Pattern Atlas* [LB21]. Mehr Informationen dazu in Kapitel 6.

4 Pattern-Sprachen für die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen

In den folgenden Abschnitten werden fünf Pattern-Sprachen vorgestellt, die im Rahmen der MLR gefunden wurden und relevante Gebiete für das Thema dieser Arbeit abdecken. In Abschnitt 4.1 werden *Cloud Computing Patterns* vorgestellt, die grundlegende Konzepte und Funktionalitäten zum Thema Cloud Computing beschreiben. Danach werden in Abschnitt 4.2 *Entperise Integration Patterns* präsentiert, die Lösungen für die Kommunikation und Integration von Anwendungen bzw. Komponenten darstellen. In Abschnitt 4.3 werden *Data Security Patterns* vorgestellt, die grundlegende Datensicherheitskonzepte beschreiben. Danach werden in Abschnitt 4.4 *Cloud Security Patterns* präsentiert, die verschiedene Sicherheitskonzepte in der Cloudumgebung darstellen. In Abschnitt 4.5 wird die Pattern-Sprache *Cloud Data Patterns for Confidentiality* vorgestellt, die verschiedene Möglichkeiten für die Speicherung von sicherheitskritischen Daten in der Cloud aufzeigen.

4.1 Cloud Computing Patterns

Im Buch *Cloud Computing Patterns* von Fehling und Kollegen [FLR+14] werden grundlegende Cloud Konzepte präsentiert, die etablierte Lösungen für wiederkehrende Probleme im Cloudumfeld aufzeigen. Dabei wurden diese Patterns in fünf Themengebiete unterteilt. In *Cloud Computing Fundamentals* werden die von Mell und Grance in *The NIST Definition of Cloud Computing* [MG+11] vorgestellten Cloud Definition analog als Patterns dargestellt. In *Cloud Offerings* werden funktionale Angebote der Cloud Anbieter präsentiert. In *Cloud Application Architectures* werden verschiedene Strukturen von Cloudanwendungen vorgestellt. In *Cloud Application Management* wird beschrieben, wie Cloudanwendungen gemanagt werden können, und in *Composite Cloud Applications* werden häufige Kombinationen von Patterns für verschiedene Anwendungsfälle aufgelistet.

Diese Pattern-Sprache ist für den Betrieb von sicherheitskritischen Daten in der Cloud relevant, da es grundlegende Technologien und Funktionalitäten der Cloud Anbieter auflistet. Diese bilden die Grundlage für den Betrieb von Daten in der Cloud und damit auch die Grundlage für Sicherheitspatterns in der Cloud.

4.2 Enterprise Integration Patterns

Das Buch *Enterprise Integration Patterns: Designing, building and deploying messaging solutions* von Hohpe und Kollegen [HW04] beschäftigt sich mit dem Thema Kommunikation zwischen verschiedenen Anwendungen oder Komponenten. Bei der Entwicklung und Integration von Anwendungen in bestehende Systeme kommt es immer wieder zu Problemen wie Asynchronität, inkompatible Datenstrukturen und vieler mehr. Aus diesem Grund bietet das Buch in seiner Pattern-Sprache 65 Patterns, die sich mit vielen dieser Probleme beschäftigen und bewiesene allgemeine Lösungen anbieten.

Diese Pattern-Sprache ist für den Betrieb von sicherheitskritischen Daten in der Cloud relevant, da sie grundlegende Methoden für die Kommunikation von verschiedenen Anwendungen miteinander präsentiert und damit die Grundlage für die Kommunikation von modernen Anwendungen ist. Aus diesem Grund basieren Sicherheitspatterns für Kommunikation auf diesen Patterns und sind somit auch für diese Arbeit relevant.

4.3 Data Security Patterns

Die Pattern-Sprache *Data Security Patterns* erstellt von Russel [Rus17] enthält insgesamt 33 Pattern zum Thema Datensicherheit. Dabei konzentrieren sich die Patterns auf die Bereiche *Confidentiality* und *Integrity*.

Diese Pattern-Sprache ist grundlegend für den Betrieb von sicherheitskritischen Daten auch im nicht Cloud-Umfeld relevant und bietet daher auch eine gute Grundlage für die Verarbeitung in der Cloud. Da sie generelle Praktiken im Bereich Datensicherheit abbilden, sind sie Grundlage für viele Patterns für Datensicherheit in der Cloud oder lassen sich auch im Cloud-Umfeld umsetzen und sind aus diesem Grund relevant für diese Arbeit.

4.4 Cloud Security Patterns

Die Pattern-Sprache *Cloud Security Patterns* erstellt und zur Verfügung gestellt vom Unternehmen *sirris* [sir] zeigt 31 Pattern zum Thema Sicherheit in der Cloud. Bei der Pattern-Sprache handelt es sich um eine Zusammenstellung verschiedenster bewiesener Lösungen, die in die folgenden fünf Bereiche unterteilt werden: *Compliance and Regulatory*, *Identification, Authentication and Authorisation*, *Secure Development, Operation and Administration*, *Privacy and Confidentiality* und *Secure Architecture*.

Diese Pattern-Sprache ist relevant, da sie Praktiken für sichere Cloud-Lösungen vorstellt und damit auch direkt aufzeigt, wie sicherheitskritische Daten in der Cloud verarbeitet werden können.

Art von sicherheitskritischen Daten	Pattern-Sprache
Authentifizierungs- und Autorisierungsdaten	Cloud Security Patterns, Data Security Patterns
Geschäftskritische Daten	Cloud Data Patterns for Confidentiality, Cloud Security Patterns, Data Security Patterns
Personenbezogene/personenbeziehbare Daten	Cloud Security Patterns

Tabelle 4.1: Abbildung der Arten von sicherheitskritischen Daten auf die vorgestellten Pattern-Sprachen

4.5 Cloud Data Patterns for Confidentiality

Die Pattern-Sprache *Cloud Data Patterns for Confidentiality* von Strauch und Kollegen [SBK+12] stellt insgesamt fünf Patterns für den Umgang und die Speicherung von sicherheitskritischen Daten in der Cloud vor. Diese Patterns decken verschiedene Szenarien und Architekturen ab, um auch zu zeigen, wie sicherheitskritische Daten für die Verarbeitung durch *Public Clouds* vorbereitet werden können.

Diese Pattern-Sprache zeigt genau, wie sicherheitskritische Daten in der Cloud und auch besonders in der *Public Cloud* verarbeitet werden können und ist daher sehr relevant für diese Arbeit. Diese Pattern-Sprache bietet auch eine Grundlage für die *Data Protection Pattern Language*.

4.6 Abbildung der Arten von sicherheitskritischen Daten auf Pattern-Sprachen

In diesem Abschnitt wird überprüft, welche Arten von sicherheitskritischen Daten durch die zuvor vorgestellen Pattern-Sprachen abgebildet werden. Die Einteilung der sicherheitskritischen Daten wurde zuvor in Abschnitt 3.2.5 durch Frage 2 eingeführt.

Der Umgang mit *Authentifizierungs- und Autorisierungsdaten* wird in den Pattern-Sprachen *Cloud Security Patterns* und *Data Security Patterns* behandelt, wie in Tabelle 4.1 zu sehen ist. Welche Patterns der jeweiligen Pattern-Sprachen sich mit dieser Art von Daten beschäftigen, ist der folgenden Auflistung zu entnehmen:

- Cloud Security Patterns
  - Multi-Factor Authentication
  - Access Token
  - Identity and Access Manager
  - Federation
- Data Security Patterns
  - Account Management

  - Access Enforcement
  - Use Of Cryptography
  - User Identification And Authentication

Wie in Tabelle 4.1 zu sehen ist, werden *geschäftskritische Daten* in den Pattern-Sprachen *Cloud Data Patterns for Confidentiality*, *Cloud Security Patterns* und *Data Security Patterns* abgebildet. In der folgenden Auflistung ist zu sehen, welche Pattern der jeweiligen Pattern-Sprachen sich mit dieser Art von Daten beschäftigen:

- Cloud Data Patterns for Confidentiality
  - Confidentiality Level Data Aggregator
  - Confidentiality Level Data Splitter
  - Filter of Critical Data
  - Pseudonymizer of Critical Data
  - Anonymizer of Critical Data
- Cloud Security Patterns
  - End-to-End Security
  - Computation on Encrypted Data
  - Data Anonymisation
  - Processing Purpose Control
  - Bastion Server
  - Automated Threat Detection
  - Virtual Network
- Data Security Patterns
  - Use Of Cryptography
  - Continuous Monitoring
  - Automated Labeling
  - Media Storage
  - Information Input Restrictions
  - Information Accuracy, Completeness, Validity, And Authenticity
  - Information Leakage
  - Physical Access Control
  - Information System Backup
  - Transmission Integrity

  - Transmission Confidentiality
  - Media Access
  - Media Labeling

In Tabelle 4.1 ist zu sehen, dass *personenbezogene/personenbeziehbare Daten* in der Pattern-Sprache *Cloud Security Patterns* behandelt wird. Dort beschäftigen sich die folgenden Patterns mit diesem Thema:

- Cloud Security Patterns
  - Data Citizenship
  - Cryptographic Erasure
  - Shared Responsibility Model
  - Compliant Data Transfer
  - Data Retention
  - Data Lifecycle
  - Intentional Data Remanence

Die meisten dieser Patterns beschäftigen sich mit bestimmten Regelungen die durch Datenschutzgesetze gerechtfertigt sind, wie beispielsweise die endgültige Löschung von Daten oder wie lange personenbezogene Daten gespeichert werden düfen.

Ganz speziell im Pattern *Data Citizenship* wird der Betrieb von personenbezogenen Daten innerhalb der EU behandelt. An dieser Stelle unterscheidet sich das Pattern von dem was durch die MLR festgestellt werden konnte. Das Pattern beschreibt, dass Cloud-Anbieter, die nicht ausschließlich innerhalb der EU tatig sind, lediglich garantieren müssen, dass die Daten die EU geographisch nicht verlassen. Wie allerdings in Abschnitt 3.6.2 festgestellt wurde reicht diese Garantie nach dem aktuellen rechtlichen Stand nicht mehr aus um datenschutzkonform zu sein. Aus diesem Grund muss das Thema Betrieb von personenbezogenen Daten in der Cloud neu definiert werden. Zusätzlich konnte im Umfang der MLR kein Pattern gefunden werden, das personenbezogene Daten selbst und die Folgen auf deren Betrieb definiert. Aus diesem Grund wurde in Kapitel 5 die Pattern-Sprache *Data Protection Pattern Language* entwickelt. Diese Pattern-Sprache definiert einerseits personenbezogene Daten als Pattern und bringt *Data Citizenship* mit Hilfe von zwei Patterns auf den aktuellen rechtlichen Stand.

5 Data Protection Pattern Language

In diesem Kapitel werden drei verschiedene Patterns vorgestellt, die sich aus den Ergebnissen der zweiten Forschungsfrage, vorgestellt in Abschnitt 3.6.2, ergeben haben. Diese Pattern-Sprache wird benötigt, da die Pattern-Sprachen, die im Zusammenhang der MLR gefunden wurden, diese Aspekte nicht abbilden. Die verschiedenen Pattern liefern dabei einen generellen Überblick über die rechtliche Lage von personenbezogenen Daten innerhalb der EU und verweisen bei konkreten Lösungen auf Patterns der Pattern-Sprachen, die in Kapitel 4 vorgestellt wurden. Es ist zu beachten, dass die vorgestellte Pattern-Sprache nicht vollständig ist und nur grundlegende Vorschriften der EU für die Speicherung und Verarbeitung von personenbezogenen Daten betrachtet. Dies liegt daran, dass ein Ziel dieser Arbeit ist eine generelle Übersicht über die Rechtslage zu schaffen. Es können allerdings noch weitere Spezialfälle existieren, die hier nicht abgebildet sind.

In Abbildung 5.1 ist eine Erweiterung des in Abschnitt 3.6.2 vorgestellten Übersichtsdiagramm zu sehen. Dieses Diagramm strukturiert die Ergebnisse der MLR zum Thema gesetzliche Regelungen. Rechtlich lassen sich Daten vereinfacht in personenbezogene und nicht personenbezogene Daten aufteilen. Für nicht personenbezogene Daten, unter die zum Beispiel geschäftskritische Daten fallen,

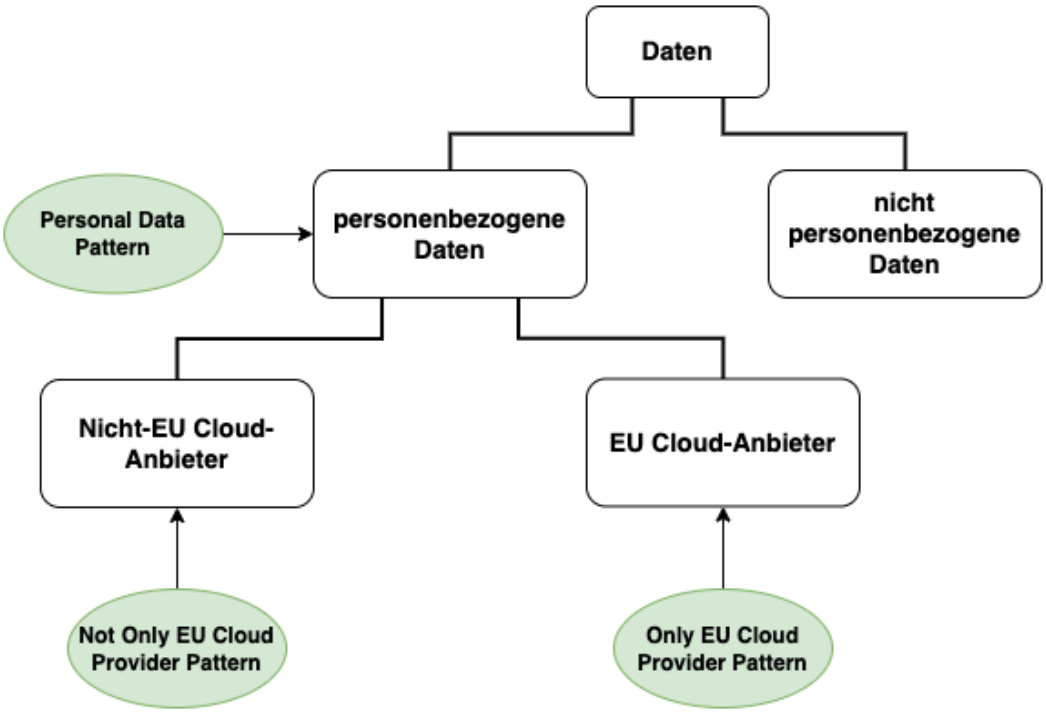

Abbildung 5.1: Visuelle Darstellung der verschiedenen rechtlichen Kategorien und ihrer Patterns

bedürfen rein rechtlich keinen besonderen Schutz, was allerdings nicht bedeutet, dass diese nicht auch schützenswert sind. Dadurch, dass hier keine besonderen rechtlichen Umstände vorliegen, ergibt sich hier allerdings kein Pattern. Personenbezogene Daten bedürfen laut EU Regelungen besonderen Schutz. Definitionen und Lösungen zu diesem Thema werden in Abschnitt 5.1 vorgestellt. Für personenbezogene Daten ergab die MLR zusätzlich noch eine weitere Unterscheidung, wenn es zum Thema Cloud kommt. Hier gibt die DSGVO vor, dass EU und nicht EU Cloud-Anbieter unterschiedlich zu behandeln sind. Aus diesem Grund ergibt sich jeweils ein Pattern für den Umgang mit den Daten in den beiden Umgebungen, die in Abschnitt 5.2 und Abschnitt 5.3 vorgestellt werden.

5.1 Personal Data Pattern

Das folgende Pattern definiert was personenbezogene Daten sind und beschreibt welche Vorkehrungen für diese Daten im Betrieb getroffen werden müssen.

Kontext

Bei personenbezogenen Daten handelt es sich um Informationen, die sich auf eine identifizierte oder identifizierbare lebende Person beziehen lassen. Auch Teilinformationen, die zur Identifizierung einer Person führen können, gelten als personenbezogene Daten. Anonymisierungen, Verschlüsselungen und Pseudonymisierungen, mithilfe derer die Person weiterhin identifiziert werden kann, sind weiterhin personenbezogene Daten. Nur Daten, die so anonymisiert wurden, dass die Person nicht mehr identifiziert werden kann, gelten nicht mehr als personenbezogene Daten. Neben den verschiedenen Regelungen, die Datenschutzgesetze zum Thema Rechte der betroffenen Personen oder Übermittlung an Drittländer vorschreiben, konzentriert sich dieses Pattern auf die Aufbewahrung von personenbezogenen Daten.

Problem

Welche Sicherheitsmaßnahmen müssen für die Aufbewahrung von personenbezogenen Daten getroffen werden, um diese entsprechend von Datenschutzgesetzen zu verwahren?

Motivation

Strafen für das Nicht-Einhalten von Datenschutzgesetzen, wie beispielsweise der DSGVO beim Umgang mit personenbezogenen Daten sind sehr hoch und können für beträchtlichen finanziellen Schaden sorgen.

Lösung

Datenschutzgesetze schreiben ein angemessenes Schutzniveau für personenbezogene Daten vor. Daten müssen daher mit ausreichend technischen Methoden geschützt werden um jeglichen unauthorisierten Zugang oder Verlust zu verhindern. Durch die Fomulierung der Vorschriften sind viele technische und organisatorische Lösungen möglich, wie beispielsweise:

- Zertifizierung des Service-Anbieters
- Verschlüsselter Datentransfer
- Implementierung nach aktuellen Sicherheitsstandards
- Verschlüsselung von personenbezogener Daten
- Regelmäßige Überprüfung, Bewertung und Evaluierung der Wirksamkeit der Sicherheitsmaßnahmen

Rule of Three

Für die Erstellung dieses Patterns wurde die Veröffentlichung der DSGVO des Europäischen Parlaments [Par16], der *California Consumer Privacy Act* [18a] und der *Data protection act* [18b] verwendet.

Verwandte Patterns

Im Folgenden werden Patterns der bereits beschriebenen Pattern-Sprachen aufgelistet, die mit diesem Pattern zusammenhängen:

- Data Security Patterns
  - Separation Of Duties
  - Least Privilege
  - Security Awareness
  - Security Training
  - Continuous Monitoring
  - Information System Backup
  - Media Access
  - Media Storage
  - Media Transport
  - Media Sanitization And Disposal
  - Physical Access Control
  - Information Leakage

- Security Categorization
- Risk Assessment
- Risk Assessment Update
- Boundary Protection
- Transmission Integrity
- Transmission Confidentiality
- Use Of Cryptography
- Information Input Restrictions
- Information Accuracy, Completeness, Validity, And Authenticity
- Information Output Handling And Retention

5.2 Only EU Cloud Provider Pattern

Das folgende Pattern präsentiert Lösungen um datenschutzkonform Daten bei einem Cloud-Anbieter zu betreiben der nur innerhalb der EU tätig ist.

Kontext

Durch das Verschieben von Anwendungen in die Cloud kommt es im Zusammenhang mit personenbezogenen Daten zu neuen Herausforderungen. Dadurch, dass die Daten nicht mehr beim Eigentümer direkt betrieben werden, sondern extern bei einem Cloud-Provider, gibt es zusätzliche Kriterien und Anforderungen, die es zu beachten gibt, um weiterhin datenschutzkonform zu sein. Da die DSGVO die Übermittlung personenbezogener Daten an Länder und Unternehmen außerhalb der EU extra regelt, gilt es diese gesondert zu behandeln.

Problem

Wie kann ich eine Anwendung bei einem EU Cloud-Anbieter betreiben und weiterhin datenschutzkonform sein?

Motivation

Die Cloud bietet durch ihre attraktive Preisgestaltung durch das Prinzip Pay-per-use viele Möglichkeiten für Einsparungen und besseren Umgang mit Lastspitzen. Auf der anderen Seite gilt allerdings innerhalb der EU die DSGVO, die viele Sicherheitsanforderungen mit sich bringt und bei nicht Einhaltung zu hohen Strafen führen kann.

Lösung

Um bei einem EU Cloud-Anbieter weiterhin datenschutzkonform zu sein, gilt es zwei Dinge zu beachten: *Zertifizierung* und *Sicherheit*.

Beim Thema Zertifizierung ist darauf zu achten, dass der Cloud-Anbieter von einer staatlich anerkannten Zertifizierungsstelle zertifiziert wurde. Es ist zu prüfen, dass diese Zertifizierung aktuell ist. Es ist auch regelmäßig zu prüfen, dass diese Zertifizierung regelmäßig wiedererteilt wird.

Bei Thema Sicherheit gilt es auch in der Cloud weiterhin ausreichenden Schutz für die Daten zu gewährleisten. Das bedeutet, dass die Verfahren, beim Betrieb im eigenen geschlossenen Netzwerk erweitert werden müssen, sodass Endpunkte oder Oberflächen im öffentlichen Umfeld der Cloud vor unerlaubtem Zugriff geschützt sind.

Rule of Three

Für die Erstellung dieses Patterns wurde die DSGVO Art. 42 [22f], der Artikel *Datenschutz in der Cloud* von Datenschutz.org [22a] und der Vortrag *Zertifizierung von Cloud-Diensten* von Maier-Reinhardt und Kollegen [MOL22] verwendet.

Verwandte Patterns

Im Folgenden werden Patterns der bereits beschriebenen Pattern-Sprachen aufgelistet, die mit diesem Pattern zusammenhängen:

- Data Protection Pattern Language
  - Personal Data Pattern
- Cloud Security Patterns
  - Virtual Network
  - Web Application Firewall
  - Certificate and Key Manager
  - Hardware Security Module
  - Secure Element
  - Secure Auditing
  - Multi-Factor Authentication
  - Federation
  - Identity and Access Manager
  - Access Token
  - Mutual Authentication

  - Per-request Authentication
  - Access Control Clearance
  - Bastion Server
  - Durable Availability
  - Automated Threat Detection
  - Vulnerability Management
- Cloud Computing Patterns
  - Private Cloud
  - Restricted Data Access Component
  - Application Component Proxy
- Enterprise Integration Patterns
  - Control Bus
  - Messaging Gateway

5.3 Not Only EU Cloud Provider Pattern

Das folgende Pattern präsentiert Lösungen um datenschutzkonform Daten bei einem Cloud-Anbieter zu betreiben der nicht nur innerhalb der EU tätig ist.

Kontext

Durch das Verschieben von Anwendungen in die Cloud kommt es im Zusammenhang mit personenbezogenen Daten zu neuen Herausforderungen. Dadurch, dass die Daten nicht mehr beim Eigentümer direkt betrieben werden, sondern extern bei einem Cloud-Provider, gibt es zusätzliche Kriterien und Anforderungen, die es zu beachten gilt, um weiterhin datenschutzkonform zu sein. Da die DSGVO auch die Übermittlung personenbezogener Daten an Länder und Unternehmen außerhalb der EU regelt, sind diese gesondert zu behandeln. Denn laut einem Urteil des EUGH muss bei der Übermittlung von personenbezogenen Daten an das EU-Ausland gewährleistet werden, dass ausländische Behörden zu keinem Zeitpunkt Zugriff auf diese Daten erhalten dürfen.

Problem

Wie kann ich eine Anwendung bei einem nicht EU Cloud-Anbieter betreiben und weiterhin datenschutzkonform sein?

Motivation

Anbieter wie AWS, GC oder MSA bieten durch ihr vielseitiges Serviceangebot viele Möglichkeiten zur Datenverarbeitung und Datenauswertung und sind dadurch für viele interessant. Dadurch, dass hierdurch allerdings Daten an ein Unternehmen mit Hauptsitz im EU-Ausland übertragen werden, ist das Verfahren besonders kritisch betreffend Datenschutz.

Lösung

Eine Lösung, weiterhin datenschutzkonform zu sein, ist nur Daten an nicht EU Cloud-Anbieter zu übermitteln, die nicht mehr als personenbezogene Daten gelten. Das bedeutet, dass diese vollständig und unumkehrbar anonymisiert wurden und anschließend dann für Auswertungszwecke verarbeitet werden.

Eine weitere Lösung wäre mithilfe von technischen Lösungen wie Verschlüsselungen, jeglichen Zugriff von ausländischen Behörden auf personenbezogene Daten zu jedem Zeitpunkt zu verhindern. Dabei ist allerdings zu beachten, dass die Verantwortung hier beim Eigentümer der Daten liegt und technisch einige Risiken bestehen.

Rule of Three

Für die Erstellung dieses Patterns wurde die Ausarbeitung *DSGVO und Nutzung US-amerikanischer Cloud-Dienste* der wissenschaftlichen Dienste [Die21], der Artikel *Internationaler datentransfer - praktische Auswirkungen der Rechtsprechung des Eugh auf den Internationalen Datentransfer (Rechtssache C-311/18 „Schrems II")* des Bundesbeauftragten für den Datenschutz und die Informationssicherheit [20] und der Artikel *Sind die schülerdaten in der Cloud Sicher?* [19c] berücksichtigt.

Verwandte Patterns

Im Folgenden werden Patterns der bereits beschriebenen Pattern-Sprachen aufgelistet, die mit diesem Pattern zusammenhängen:

- Data Protection Pattern Language
  - Personal Data Pattern
  - Only EU Cloud Provider Pattern
- Cloud Data Patterns for Confidentiality
  - Confidentiality Level Data Aggregator
  - Confidentiality Level Data Splitter
  - Filter of Critical Data
  - Pseudonymizer of Critical Data

  - Anonymizer of Critical Data
- Cloud Security Pattern
  - Data Citizenship
  - Shared Responsibility Model
  - Compliant Data Transfer
- Enterprise Integration Patterns
  - Content Filter

6 Decision Support System

In den folgenden Abschnitten wird ein Decision Support System beschrieben, das einen Nutzer dabei unterstützen soll, sicherheitskritische Daten in der Cloud sicher verarbeiten zu können. In Abschnitt 6.1 werden die Anforderungen an das Decision Support System erläutert, die vorher mithilfe von Experteninterviews ermittelt wurden. Anschließend wird in Abschnitt 6.3 die Datenstruktur dieses Systems dargestellt und erklärt. Zum Schluss wird in Abschnitt 6.4 der Prototyp, der im Rahmen dieser Arbeit umgesetzt wurde, vorgestellt und erklärt, wie dort die vorher dargestellten Anforderungen eingebunden wurden.

6.1 Anforderungen

In Abschnitt 3.2.5 wurden die Ergebnisse des Interviews vorgestellt, das im Rahmen dieser Arbeit durchgeführt wurde. Eines der Ziele des Interviews war es, Anforderungen für ein Decision Support System zu ermitteln, das einem Nutzer dabei helfen soll, sicherheitskritische Daten in der Cloud betreiben bzw. verarbeiten zu können. Im Folgenden werden die Anforderungen aufgelistet und jeweils kurz beschrieben.

- Generelle Übersicht

  Die Experten beschrieben bei dieser Anforderung eine generelle Übersichtsseite, die dem Nutzer einen einfachen Einstieg in das Thema bietet. Es sollten also übersichtlich alle Informationen, die das System bietet, dargestellt werden und auf einen Blick erkennbar sein, was den Nutzer erwartet.

- Best Practices / Patterns

  Die Experten führten im Interview aus, dass verschiedene Best Practices bzw. Pattern zu dem Thema übersichtlich aufgelistet werden sollten. Die Struktur, die Pattern vorgeben, eignet sich besonders gut für eine einheitliche Dokumentation (die Struktur von Pattern ist in Abschnitt 2.2 zu finden). Diese einheitliche Dokumentation ist auch für ein gutes Verständnis hilfreich.

- Kategorisierung

  Die verschiedenen Best Practices bzw. Pattern sollten gut kategorisiert sein, um eine einfache Suche und Einordnung zu bieten.

- Tags

  Das Hinzufügen von Schlagwörtern ermöglicht dabei noch zusätzliche Möglichkeiten der Einordnung des Patterns zu seinem Themengebiet.

- Verlinkungen zwischen Pattern

  Die Experten wiesen darauf hin, dass Verknüpfungen zwischen den Informationen hilfreich sind. Analog zu einer Pattern-Sprache, wie in Abschnitt 2.3 beschrieben. Dadurch ist es deutlich einfacher, tiefes Verständnis für das Thema zu entwickeln.

- Verlinkung zu konkreten Umsetzungen

  Neben den Verlinkungen zwischen Lösungen stellten die Experten eine zusätzliche Verlinkung zu konkreten Umsetzungen des Konzepts dar, das durch die Lösung beschrieben wird. Dieses Konzept wurde bereits im Artikel *From Pattern Languages to Solution Implementations* von Falkenthal und Kollegen [FBB+14] für Patterns vorgestellt. Dieser Artikel beschäftigt sich mit der Problematik, die auch die Experten aufgezeigt haben. Da Patterns eine generalisierte Lösung für wiederkehrende Probleme beschreiben, gestaltet sich oft die konkrete Umsetzung schwierig bzw. muss diese noch gesucht werden. Analog zur Lösung des Artikels beschrieben die Experten die gewünschte Funktion der Verlinkung mit konkreten Umsetzungsvorschlägen abhängig der Technologie.

- Suchfunktion

  Eine weitere Anforderung ist eine Suchfunktion. Dabei soll es möglich sein, mithilfe von Schlagwörtern nach passenden Patterns zu suchen.

- Filter

  Filter sind eine weitere Anforderung. Bei diesen soll es möglich sein, anhand eines Themas oder Technologie Patterns zu filtern und dadurch nur relevante Informationen zu sehen.

6.2 Konzept

Die in Abschnitt 6.1 vorgestellten Anforderungen geben bereits einen klaren Rahmen für das Decision Support System vor. In diesem Abschnitt wird ein Konzept für eine Anwendung präsentiert, das die aufgelisteten Anforderungen umsetzt. Für die Architektur der Anwendung bietet sich eine klassische Webanwendung mit einem Web-Front-End und einem REST-Backend mit Datenbank an. Da es sich dabei um eine klassische und allgemein bekannte Architektur handelt, wird diese nicht näher beschrieben.

Laut den Anforderungen benötigt das Decision Support System eine Hauptseite, die es möglich macht, sich eine Übersicht über die enthaltenen Informationen zu verschaffen. Aus diesem Grund sollte die Anwendung eine Seite enthalten, auf der alle enthaltenen Pattern in einer vereinfachten Form darstellen. Dies kann mithilfe der Icons und der Namen der Patterns umgesetzt werden, wie beispielhaft in Abbildung 6.1 zu sehen ist. Dadurch, dass die Icons und der Name bereits einiges über die Pattern aussagen, können bereits hier Rückschlüsse auf ihren Inhalt gezogen werden. Des Weiteren sollte auf dieser Hauptseite die Such- und Filterfunktion umgesetzt werden. Bei der Suchfunktion bietet sich ein Freitextfeld und bei der Filterfunktion ein Dropdown für alle verfügbaren Filter an.

Nach der Hauptseite wird noch eine Detailseite benötigt, ein Beispiel dafür ist in Abbildung 6.1 zu sehen. Diese zeigt alle verfügbaren Informationen über ein ausgewähltes Pattern an. Hier werden alle Attribute eines Pattern wie Name, Icon, Problem etc. angezeigt (Attribute eines Pattern sind in

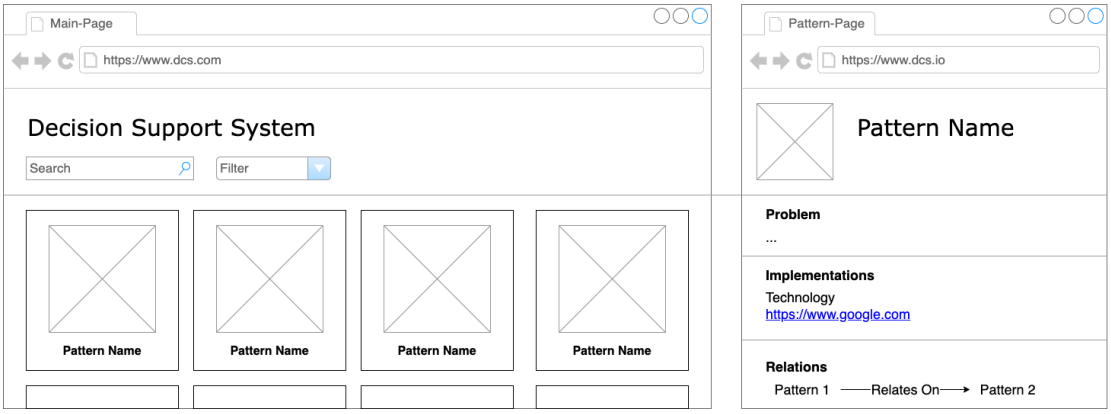

Abbildung 6.1: Visuelle Darstellung einer Umsetzungsmöglichkeit des Konzepts

Abschnitt 2.2 zu finden). Auf dieser Seite werden ebenfalls mithilfe von Verlinkungen konkrete Umsetzungsvorschläge für das dargestellte Pattern angezeigt, die mit der verwendeten Technologie gekennzeichnet sind. Ebenfalls werden die Verlinkungen zu anderen Patterns dargestellt. Hier kann beispielsweise mithilfe von beschrifteten Pfeilen die Richtung und Typ der Relation angezeigt werden.

6.3 Datenstruktur

In Abbildung 6.2 ist die Datenstruktur zu sehen, die sich aus den Anforderungen ergibt und in Abschnitt 6.1 beschrieben wurden. Durch die Verwendung von Pattern ergibt sich eine klare Struktur mit festen Attributen: *Name*, *Icon*, *Context*, *Problem*, *Forces*, *Solution*, *Sketch* und *Results*. Deren zugehörige Datentypen sind in Abbildung 6.2 zu sehen. Jedes Pattern gehört dabei zu einer Pattern-Sprache. Diese hat einen *Namen* und *Icon*. Eine Pattern-Sprache kann mehrere Pattern enthalten. Für die Anforderung der Kategorisierung erhält ein Pattern das zusätzliche Attribut *Category*. Dort wird mithilfe eines Strings die zugehörige Kategorie festgehalten. Für die Anforderung von Tags erhält das Pattern ein weiteres Attribut mit dem Namen *Tags*. Dort kann in einem String eine kommagetrennte Liste von Schlagwörtern gespeichert werden.

Die Relationen eines Patterns werden in *Relation* definiert. Eine *Relation* enthält die folgenden Attribute: *Type*, *Pattern1*, *Pattern2* und *Direction*. Der *Type* definiert die Art der Relation. Hier können beliebig viele Typen angelegt werden, wie beispielsweise *RelatesTo*. *Pattern1* und *Pattern2* geben die beiden Patterns an, die in Relation zueinanderstehen und *Direction* die Richtung der Relation. Dabei ist *Left*, *Right* oder *LeftRight* möglich. Hier wird *Pattern1* als das linke und *Pattern2* als das rechte Pattern definiert und die Richtung gibt an, auf welche Seite die Relation zeigt. Ein Pattern kann mehrere Relationen besitzen. Eine Relation ist immer zwei Pattern zugehörig. Aus diesem Grund handelt es sich hier um eine 2 zu n Beziehung.

Die Implementierungen eines Patterns sind in *Implementation* definiert. Eine *Implementation* hat die Attribute *Type* und *Link*. Der *Type* definiert die Art der Implementierung und kann frei als Text definiert werden. Hierbei kann beispielsweise eine Plattform oder Technologie, auf der die Implementierung basiert, angegeben werden. Bei dem Attribut *Link* wird der dazugehörige Link zu

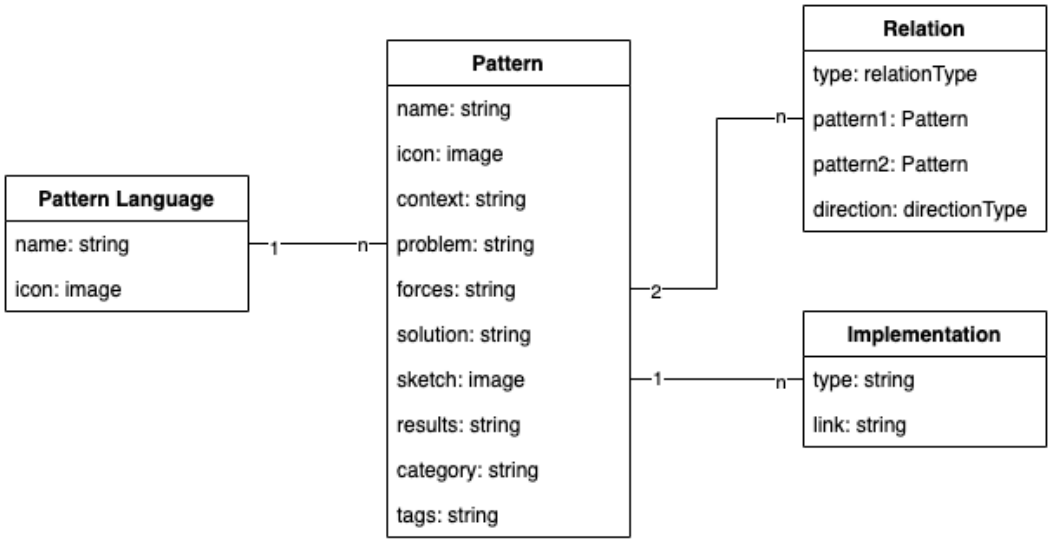

Abbildung 6.2: Datenstruktur des Decision Support Systems

Anforderung	
Generelle Übersicht	✓
Best Practices / Patterns	✓
Kategorisierung	✓
Tags	✓
Verlinkungen zwischen Pattern	✓
Verlinkungen zwischen Pattern anderer Pattern-Sprachen	✓
Verlinkung zu konkreten Umsetzungen	✓
Suchfunktion	✓
Filter	✓

Tabelle 6.1: Anforderungen für das Decision Support System und wie diese umgesetzt wurden

der Dokumentation der Implementierung angegeben. Ein Pattern kann mehrere *Implementations* besitzen, eine *Implementation* gehört aber immer nur zu einem Pattern. Dies ist durch die 1 zu n Beziehung definiert.

6.4 Umsetzung

Für die Umsetzung des Prototypen wurde der Pattern Atlas als Grundlage verwendet. Dieser wurde in einem Kapitel vom Buch *Next-Gen Digital Services. A Retrospective and Roadmap for Service Computing of the Future* von Leymann und Barzen [LB21] vorgestellt. Es wurde sich für den Pattern Atlas entschieden, da dieser bereits einige Anforderungen aus Abschnitt 6.1 umsetzt. Zudem ist die Datenstruktur des Pattern Atlas bereits ähnlich zu der in Abschnitt 6.3 vorgestellten Datenstruktur.

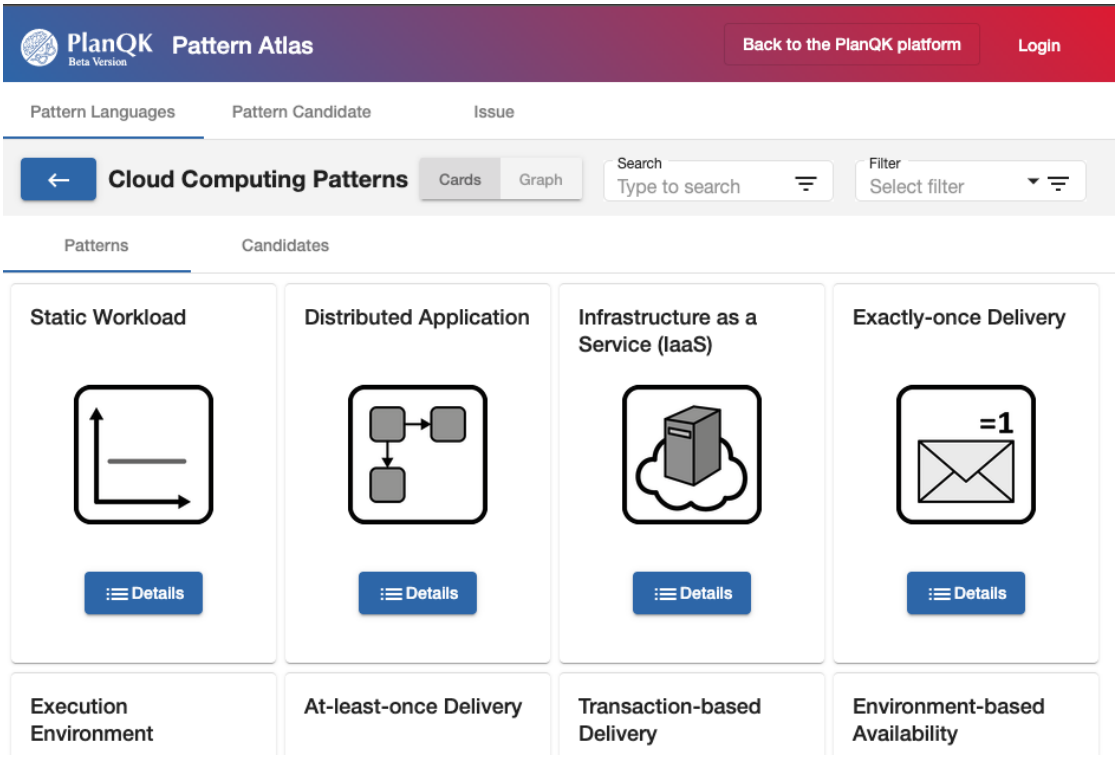

Abbildung 6.3: Screenshot der Hauptseite des Prototypen

In Tabelle 6.1 sind diese Anforderungen aufgelistet und dargestellt, wie diese umgesetzt wurden. Grün markierte Anforderungen waren bereits im Pattern Atlas enthalten. Orange markierte Anforderungen wurden im Rahmen dieser Arbeit umgesetzt.

Der Patternatlas bietet die Möglichkeit, mehrere Pattern-Sprachen übersichtlich darzustellen. Dadurch sind zunächst die Pattern in ihre Pattern-Sprache gruppiert und werden bei der Auswahl einer Pattern-Sprache entsprechend aufgelistet. Hier wird ein Pattern mithilfe des Icons und des Namens vereinfacht dargestellt. Eine Suchfunktion hat der Pattern Atlas ebenfalls bereits auf diesen Seiten. Bei der Auswahl eines Patterns werden alle Informationen über das Pattern dargestellt. Hier werden bereits Verlinkungen zwischen Pattern innerhalb einer Pattern-Sprache angezeigt. Verbindungen zu Pattern anderer Sprachen ist bisher nicht möglich.

Alle in Tabelle 6.1 Orange markierten Anforderungen wurden im Rahmen dieser Arbeit umgesetzt und sind im GitHub des Pattern Atlas unter folgendem Link zu finden *https://github.com/PatternAtlas*.

In Abbildung 6.3 ist die Hauptseite des Prototypen zu sehen. Dort wurde im Rahmen der Arbeit die Filterfunktion hinzugefügt und die Suchfunktion auf die Tags erweitert. Bei einem Klick auf eines der Pattern wird man auf die Detailseite weitergeleitet. Diese ist Auszugsweise in mehreren Screenshots zu sehen. Es wurde sich auf die Teile beschränkt, die im Rahmen dieser Arbeit hinzugefügt wurden.

In Abbildung 6.4 ist der obere Teil der Detailseite zu sehen. Hier wurde die Kategorie hinzugefügt. In diesem Beispiel ist das *NIST*. In Abbildung 6.5 ist der untere Teil der Detailseite zu sehen. Zwischen den beiden Screenshots stehen die verschiedenen Informationen, die ein Pattern enthält.

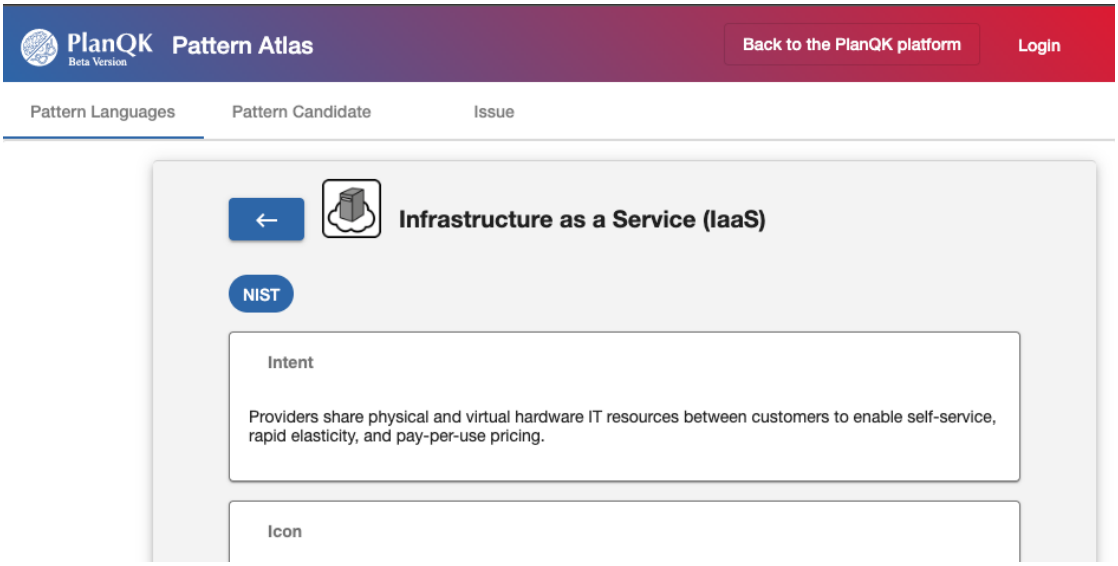

Abbildung 6.4: Screenshot des oberen Teil der Detailseite des Prototypen

Im Block *Implementations* ist die Umsetzung der Verlinkung zu sehen. Der Block *Relations to other Patterns* enthält alle Verbindungen mit anderen Patterns. Hier wurde der *Pattern Atlas* erweitert um Verbindungen zu Pattern anderer Patternsprachen zu ermöglichen. Unterhalb dieses Blocks sind die *Tags* dieses Patterns zu sehen. Diese wurden hier beispielhaft hinzugefügt.

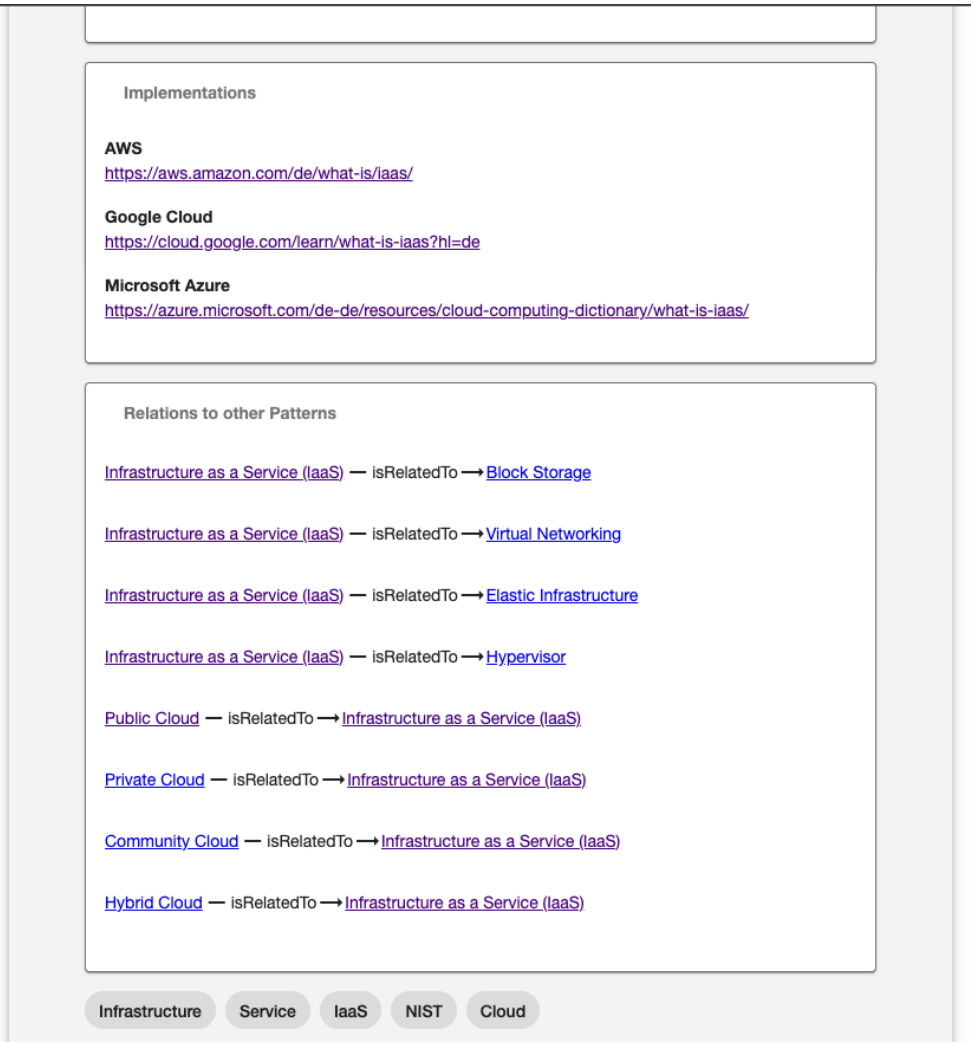

Abbildung 6.5: Screenshot des unteren Teil der Detailseite des Prototypen

7 Zusammenfassung und Ausblick

Durch die Experteninterviews konnten insgesamt drei Arten von sicherheitskritischen Daten ermittelt werden: *Personenbezogene bzw. personenbeziehbare Daten*, *geschäftskritische Daten* und *Authentifizierungs- und Autorisierungsdaten*.

Mit Hilfe der MLR konnte gezeigt werden, dass bereits einige Pattern-Sprachen existieren, die wiederkehrende Probleme bei der Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen, lösen. Es wurden insgesamt fünf Pattern-Sprachen ermittelt, die in diesem Themenbereich sehr hilfreich sein können: *Cloud Computing Patterns*, *Enterprise Integration Patterns*, *Data Security Patterns*, *Cloud Security Patterns* und *Cloud Data Patterns for Confidentiality*. Des Weiteren konnte bei der MLR ermittelt werden, dass bestimmte Aspekte von diesen Pattern-Sprachen nicht abgebildet werden. Dafür wurde im Rahmen dieser Arbeit die Pattern-Sprache *Data Protection Pattern Language* erstellt. Sie enthält drei Patterns, die den rechtlichen Aspekt von sicherheitskritischen Daten innerhalb der EU abbilden.

Das *Personal Data Pattern* bietet Lösungen für die Speicherung von personenbezogenen Daten innerhalb der EU. In der EU gelten besondere Datenschutzrichtlinien, die einen besonderen Umgang mit diesen Daten vorschreibt. Dadurch gibt es einige Dinge zu beachten, die dieses Pattern beschreibt, wenn man diese Art von Daten speichern und verarbeiten möchte. Des Weiteren wird in diesem Pattern auch dargestellt, wie die Daten definiert sind, die in den Geltungsbereich der DSGVO fallen.

Das *Only EU Cloud Provider Pattern* beschreibt, welche zusätzlichen Dinge zu beachten sind, wenn man die im *Personal Data Pattern* definierten Daten in einer Cloud eines EU Cloud-Provider betreiben möchte.

Das *Not Only EU Cloud Provider Pattern* zeigt, was zu beachten ist, wenn man personenbezogene Daten bei einem Cloud-Anbieter außerhalb der EU betreiben möchte. Diese Unterscheidung kommt dadurch, dass nach der DSGVO, EU und Nicht-EU Anbieter unterschiedlich zu behandeln sind.

Durch die Experteninterviews, die in dieser Arbeit durchgeführt wurden, war es möglich, detaillierte Anforderungen für ein Decision Support System zu sammeln, die den Nutzer dabei unterstützt, sicherheitskritische Daten in der Cloud verarbeiten zu können. Diese wurden in den Forschungsprototypen *Pattern Atlas* mit integriert. Der *Pattern Atlas* bereits viele Grundfunktionen um, die sich aus den Anforderungen ergeben haben. Zusätzlich war die zugrunde liegende Datenstruktur bereits sehr ähnlich. Daraufhin wurde der *Pattern Atlas* um die angeforderten Funktionen erweitert. Dies beinhaltete beispielsweise Funktionen wie: Verknüpfung eines Patterns mit konkreten Umsetzungsvorschlägen, Relationen mit Pattern, einer anderen Pattern-Sprache, Tags und Kategorien.

Insgesamt wurden also alle vorher definierten Ziele dieser Arbeit erfüllt. Es wurden Pattern-Sprachen für die Verarbeitung von sicherheitskritischen Daten durch Cloud-Anwendungen gefunden, dokumentiert und miteinander verbunden. Zudem konnte mithilfe eines Prototypen gezeigt werden, wie diese Informationen einfacher zugänglich gemacht werden können.

Ausblick

Weiterführende Forschungen könnten im Bereich der Pattern-Sprache *Data Protection Pattern Language* betrieben werden, die im Rahmen dieser Arbeit entstanden ist. Wie bereits in der Arbeit hingewiesen wurde, ist diese Sprache nicht vollständig, sondern bildet drei Grundlagen-Patterns ab. Diese bilden allgemeine Regelungen ab, die beim Betrieb von personenbezogenen Daten in der Cloud zu beachten sind. Dadurch liegt hier noch weiteres Forschungspotenzial. Es könnte geprüft werden, welche technische Lösungen für eine bessere Verwendung von Cloud-Anbietern, die nicht nur innerhalb der EU tätig sind, existieren und in Form weiterer Patterns festgehalten werden. Zusätzlich könnten rechtliche Spezialfälle und deren technische Herausforderungen geprüft und abgebildet werden.

Literaturverzeichnis

[17] *10 Business Data Security Tips to Protect Your Company: C1c*. Jan. 2017. URL: https://www.c1c.net/blog/business-data-security-tips-protect-company/ (zitiert auf S. 36–38).

[18a] *California Consumer Privacy Act*. 2018. URL: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5 (zitiert auf S. 49).

[18b] *Data protection act 2018*. 2018. URL: https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted (zitiert auf S. 49).

[19a] Sep. 2019. URL: https://datadome.co/bot-management-protection/web-application-security-best-practices/ (zitiert auf S. 36, 37).

[19b] *Gesetz zum Schutz von Geschäftsgeheimnissen (geschgehg)*. Apr. 2019. URL: https://www.gesetze-im-internet.de/geschgehg/BJNR046610019.html (zitiert auf S. 38).

[19c] *Sind die schülerdaten in der Cloud Sicher?* Nov. 2019. URL: https://www.treffpunkt-kommune.de/sind-die-schuelerdaten-in-der-cloud-sicher/ (zitiert auf S. 53).

[19d] *Was Sind Personenbezogene Daten?* Nov. 2019. URL: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_de (zitiert auf S. 39).

[20] *Internationaler datentransfer - praktische Auswirkungen der Rechtsprechung des Eugh auf den Internationalen Datentransfer (Rechtssache C-311/18 „Schrems II")*. 2020. URL: https://www.bfdi.bund.de/DE/Fachthemen/Inhalte/Europa-Internationales/Auswirkungen-Schrems-II-Urteil.html (zitiert auf S. 20, 53).

[21] *Everything you need to know about... protecting your customers' data*. Juni 2021. URL: https://mailchimp.com/courier/article/protecting-customer-data/ (zitiert auf S. 38).

[22a] Aug. 2022. URL: https://www.datenschutz.org/cloud/ (zitiert auf S. 51).

[22b] *15 application security best practices 2022*. Sep. 2022. URL: https://snyk.io/learn/application-security-best-practices/ (zitiert auf S. 36, 37).

[22c] *6 tips to protect your company's data*. Juli 2022. URL: https://www.bdc.ca/en/articles-tools/technology/invest-technology/tips-protect-company-data (zitiert auf S. 36–38).

[22d] *A step-by-step guide to cloud security best practices*. 2022. URL: https://www.skyhighsecurity.com/en-us/cybersecurity-defined/cloud-security-best-practices.html (zitiert auf S. 37).

[22e] *Data Protection and Privacy: 12 ways to protect user data*. Juni 2022. URL: https://cloudian.com/guides/data-protection/data-protection-and-privacy-7-ways-to-protect-user-data/ (zitiert auf S. 36–38).

[22f] *Datenschutz-Grundverordnung: DSGVO Als übersichtliche Seite*. Sep. 2022. URL: https://dsgvo-gesetz.de/ (zitiert auf S. 51).

[Ale77] C. Alexander. *A pattern language: towns, buildings, construction*. Oxford university press, 1977 (zitiert auf S. 15, 16).

[Bit21] L. Biton. *Secure authentication - everything you need to know*. Nov. 2021. URL: https://www.securecoding.com/blog/secure-authentication/ (zitiert auf S. 37).

[Che22] S. Cherednichenko. *11 web application security best practices you need to know*. Sep. 2022. URL: https://www.mobindustry.net/blog/11-web-application-security-best-practices-you-need-to-know/ (zitiert auf S. 36, 37).

[Coo22] A. Coos. *5 ways large enterprises protect their data*. Juni 2022. URL: https://www.endpointprotector.com/blog/5-ways-big-companies-protect-their-data/ (zitiert auf S. 37).

[Die21] W. Dienst. *DSGVO und Nutzung US-amerikanischer Cloud-Dienste*. 2021 (zitiert auf S. 19, 39, 53).

[FBB+14] M. Falkenthal, J. Barzen, U. Breitenbücher, C. Fehling, F. Leymann. „From Pattern Languages to Solution Implementations“. Englisch. In: *Proceedings of the Sixth International Conferences on Pervasive Patterns and Applications (PATTERNS 2014)*. Xpert Publishing Services, Mai 2014, S. 12–21. ISBN: 978-1-61208-343-8. URL: http://www2.informatik.uni-stuttgart.de/cgi-bin/NCSTRL/NCSTRL_view.pl?id=INPROC-2014-37&engl= (zitiert auf S. 56).

[FBL18] M. Falkenthal, U. Breitenbücher, F. Leymann. „The nature of pattern languages“. In: *Pursuit of pattern languages for societal change* (2018), S. 19 (zitiert auf S. 16, 17).

[FLR+14] C. Fehling, F. Leymann, R. Retter, W. Schupeck, P. Arbitter. *Cloud computing patterns: fundamentals to design, build, and manage cloud applications*. Springer, 2014 (zitiert auf S. 41).

[GFM19] V. Garousi, M. Felderer, M. V. Mäntylä. „Guidelines for including grey literature and conducting multivocal literature reviews in software engineering“. In: *Information and Software Technology* 106 (2019), S. 101–121 (zitiert auf S. 19, 20, 34, 35).

[Har22] C. Harvey. *Top 12 cloud security best practices [2022]: Esecurity planet*. Juni 2022. URL: https://www.esecurityplanet.com/cloud/cloud-security-best-practices/ (zitiert auf S. 37).

[HRRR12] M. Host, A. Rainer, P. Runeson, B. Regnell. *Case Study Research in Software Engineering: Guidelines and Examples*. John Wiley und Sons, 2012 (zitiert auf S. 21).

[HW04] G. Hohpe, B. Woolf. *Enterprise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley Professional, 2004 (zitiert auf S. 42).

[Inv22] Investopedia. *Best Practices*. 2022. URL: https://www.investopedia.com/terms/b/best_practices.asp (zitiert auf S. 15).

[Job19] A. Jobard. *10 best practices to protect your users' data*. Feb. 2019. URL: https://medium.com/tanker-blog/10-best-practices-to-protect-your-users-data-f8fb64d46f09 (zitiert auf S. 36–38).

[Jon22] J. Jones. *11 best practices for developing secure web applications*. Feb. 2022. URL: https://www.lrswebsolutions.com/Blog/Posts/32/Website-Security/11-Best-Practices-for-Developing-Secure-Web-Applications/blog-post/ (zitiert auf S. 36, 37).

[Kri21] A. Krishna. *7 web application security practices you can use*. Sep. 2021. URL: https://www.thesslstore.com/blog/web-application-security-practices-you-can-use/ (zitiert auf S. 36, 37).

[Lam21] S. Lamatrice. *Data security: Authentication, Authorization and Encryption*. Juni 2021. URL: https://www.progress.com/blogs/data-security-basics-authentication-authorization-encryption-auditing (zitiert auf S. 36, 37).

[LB21] F. Leymann, J. Barzen. „Pattern Atlas“. In: *Next-Gen Digital Services. A Retrospective and Roadmap for Service Computing of the Future*. Hrsg. von M. Aiello, A. Bouguettaya, D. A. Tamburri, W.-J. van den Heuvel. Cham: Springer International Publishing, 2021, S. 67–76. ISBN: 978-3-030-73203-5. DOI: 10.1007/978-3-030-73203-5_5. URL: https://doi.org/10.1007/978-3-030-73203-5_5 (zitiert auf S. 39, 58).

[Lim20] N. Limbachiya. *7 web application security best practices - dzone security*. Okt. 2020. URL: https://dzone.com/articles/7-web-application-security-best-practices (zitiert auf S. 38).

[MG+11] P. Mell, T. Grance et al. „The NIST definition of cloud computing“. In: (2011) (zitiert auf S. 41).

[MOL22] D. N. Maier-Reinhardt, K. Osterhage, S. Lins. *Zertifizierung von Cloud-Diensten*. Feb. 2022. URL: https://stiftungdatenschutz.org/veranstaltungen/unsere-veranstaltungen-detailansicht/zertifizierung-von-cloud-diensten-199 (zitiert auf S. 51).

[Nid22] T. A. Nidecki. *7 web application security best practices*. Juli 2022. URL: https://www.acunetix.com/blog/web-security-zone/7-web-application-security-best-practices/ (zitiert auf S. 36, 37).

[Omo22] T. Omoth. *Ten ways to protect your company from the next big data breach*. Feb. 2022. URL: https://www.itpro.co.uk/security/data-breaches/358455/10-ways-to-protect-your-company-from-the-next-big-data-breach/ (zitiert auf S. 36, 37).

[Par16] E. Parlament. *Verordnung (EU) 2016/679 des Europäischen Parlaments und des 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)*. Mai 2016. URL: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32016R0679 (zitiert auf S. 49).

[Rah14] M. Rahal. *10 tips for entrepreneurs on how to protect users' private information*. Jan. 2014. URL: https://www.wamda.com/2014/01/10-tips-how-to-protect-users-private-information (zitiert auf S. 36, 37).

[Rus17] Russell. *SP-013: Data Security Pattern*. März 2017. URL: https://www.opensecurityarchitecture.org/cms/library/patternlandscape/259-pattern-data-security (zitiert auf S. 36–38, 42).

[SBK+12] S. Strauch, U. Breitenbuecher, O. Kopp, F. Leymann, T. Unger. „Cloud Data Patterns for Confidentiality“. In: *Proceedings of the 2nd International Conference on Cloud Computing and Service Science, CLOSER 2012, 18-21 April 2012, Porto, Portugal*. SciTePress, 2012, S. 387–394 (zitiert auf S. 43).

[Sch19] J. Schneider. *Datenschutz: nach der EU-Datenschutz-Grundverordnung*. CH Beck, 2019 (zitiert auf S. 13).

[Sel20] A. Selzer. *Datenschutzrechtliche Zulässigkeit von Cloud-Computing-Services und deren teilautomatisierte Überprüfbarkeit*. Springer, 2020 (zitiert auf S. 19).

[sir] sirris. URL: http://www.sirris.be.s3-website-eu-west-1.amazonaws.com/ (zitiert auf S. 37, 38, 42).

Alle URLs wurden zuletzt am 18. 10. 2022 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Ort, Datum, Unterschrift