



**Bachelorarbeit**

**zu dem Thema**

# Untersuchung der Klimaauswirkungen des Datenschutzes

*– On Data Protection Carbon –*

**eingereicht von**

[redacted]

[redacted]

# Vorwort und Danksagungen

*Homo homini lupus*, der Mensch sei dem Menschen ein Wolf, schrieb der Staatstheoretiker und Philosoph Thomas Hobbes. Eine gute Metapher für die Probleme, welche dieser Arbeit zugrundeliegen: das Verderben der Menschheit durch Menschenhand.

Zwei Beispiele für derartige Probleme sind der Klimawandel und der zunehmende Missbrauch unserer persönlichen Informationen durch Unternehmen und Behörden. Da mir die Bekämpfung beider dieser Probleme am Herzen liegt, freut es mich natürlich sehr, ein Thema für eine Bachelorarbeit gefunden zu haben, das die beiden Gebiete verbindet.

*Nach* Fertigstellung dieser Arbeit habe ich aus Interesse ChatGPT gefragt, was die Klimaauswirkungen des Datenschutzes seien. Mir wurde folgendes geantwortet

> „Der Datenschutz an sich hat keine direkten Klimaauswirkungen. Datenschutz bezieht sich auf die Prinzipien und Maßnahmen, die den Schutz persönlicher Daten und die Privatsphäre von Personen gewährleisten sollen. Diese Maßnahmen betreffen hauptsächlich die Art und Weise, wie Daten gesammelt, gespeichert, verarbeitet und geteilt werden, insbesondere in digitalen Umgebungen.
>
> Es gibt jedoch eine indirekte Verbindung zwischen Datenschutz und Umweltauswirkungen, die sich aus dem zunehmenden Datenverkehr und dem steigenden Energieverbrauch in Rechenzentren ergeben. Die verstärkte Nutzung von Online-Diensten, Cloud-Speicher und digitalen Plattformen bedeutet, dass mehr Daten gespeichert und verarbeitet werden müssen. Dies führt zu einem höheren Bedarf an Rechenleistung und somit zu einem erhöhten Energieverbrauch in den Rechenzentren, die diese Daten hosten. Ein hoher Energieverbrauch wiederum trägt zur Freisetzung von Treibhausgasen und damit zum Klimawandel bei.
>
> Es ist wichtig zu betonen, dass der Datenschutz an sich notwendig ist, um die Privatsphäre und die Grundrechte der Menschen zu schützen. Allerdings ist es auch wichtig, nachhaltige Praktiken in der Datenverarbeitung zu fördern, um die Umweltauswirkungen zu minimieren. Unternehmen und Organisationen können beispielsweise in energieeffiziente Rechenzentren investieren, erneuerbare Energiequellen nutzen und Daten effizienter speichern und verarbeiten, um den Energieverbrauch zu reduzieren und damit die Klimaauswirkungen zu minimieren.“

Offenbar „kennt“ ChatGPT den Begriff des technischen Datenschutzes nicht, es „verkennt“ zudem, dass Datenschutz nicht darin besteht, noch mehr personenbezogene Daten in „Online-Dienste, Cloud-Speicher und digitale Plattformen“ zu stecken.

Der Schutz von Grundrechten und Grundfreiheiten sollte daher wohl auf absehbare Zeit den Menschen selber überlassen bleiben.

Das gleiche gilt offensichtlich für die Untersuchung der Klimafolgen des Datenschutzes.

Im Zusammenhang mit der Erstellung dieser Arbeit gebührt mein Dank
– ohne besondere Reihenfolge –

1. [redacted]
2. [redacted]
3. [redacted]
4. [redacted]

[redacted]

# Inhaltsverzeichnis

# Kapitel 1

# Problemstellung

## 1.1 Hintergrund

Der Klimawandel ist, ohne dass hieran in Wissenschaft und Forschung noch ernsthafte Zweifel erhoben werden, real.[1] Er ist menschengemacht[2] und seine Auswirkungen sind bereits jetzt dramatisch.[3]

An den Ursachen des Klimawandels sind auch die Computer beteiligt [11], die für unsere moderne digitale Gesellschaft unerlässlich sind. Ein Computer benötigt zum Betrieb Strom, der in der Regel nicht oder nicht vollständig ohne Erzeugung von klimaschädlichen Emissionen beschafft werden kann. Zudem sind sowohl Herstellung als auch Entsorgung bzw. Recycling von Computern in aller Regel nicht klimaneutral, da z. B. Rohstoffe wie Seltene Erden durch Abbau gewonnen werden müssen.

Wenn man jedoch dem liberalen Feuilleton zuhört, könnte der Eindruck entstehen, dass es eine noch schlimmere Gefahr als den Klimawandel gibt, nämlich den Datenschutz, der u. a. an der schlechten Bekämpfung der Corona-Pandemie [30], am Tod von Menschen [21], an der Verunstaltung des Internets [38], am Ausbleiben des Nikolaus (sic!) [37] u. v. m. Schuld haben soll.

Von anderer Seite wird hingegen die Meinung vertreten, dass es sich beim Datenschutz bloß um ein Instrument des Schutzes von Grundrechten und Grundfreiheiten und somit um einen Garant der Menschenwürde innerhalb einer modernen freiheitlich-demokratischen Gesellschaft handelt [28, 32]. Ein solches Datenschutzverständnis erscheint, insbesondere da in der datenschutzrechtlichen Fachöffentlichkeit viele schlüssige Argumente gegen die Meinung des liberalen Feuilletons vorgebracht werden, zunächst vorzugswürdiger.

[1] „Human activities, principally through emissions of greenhouse gases, have unequivocally caused global warming“ [4].
[2] Ebenda.
[3] „Human-induced climate change, including more frequent and intense extreme events, has caused widespread adverse impacts and related losses and damages to nature and people [...]. The rise in weather and climate extremes has led to some irreversible impacts as natural and human systems are pushed beyond their ability to adapt.“ [33].

Dieses Vertrauen in den Datenschutz könnte jedoch erschüttert werden, wenn sich herausstellen würde, dass der Datenschutz tatsächlich katastrophale Auswirkungen auf die Gesellschaft hätte. Würde beispielsweise die Umsetzung von Datenschutzmaßnahmen in unangemessenem Umfang Klimanachteile bedingen, so würde sich berechtigterweise die Frage nach dem Vorrang des Klimaschutzes vor dem Datenschutz stellen.

Um diese Frage auf eine plausible Tatsachengrundlage stellen zu können, soll diese Arbeit untersuchen, ob und inwieweit Datenschutz klimaschädlich sein kann. Zudem soll diese Arbeit aufzeigen, in welcher Weise Datenschutzmaßnahmen mit CO2-Emissionen korrelieren können, in der Hoffnung, dass dieser Faktor ggf. bei der Implementierung von Datenverarbeitungssystemen zusätzlich berücksichtigt wird, sodass Systeme entwickelt werden können, die sowohl datenschutzfreundlich als auch klimafreundlich sind.

Gegenstand dieser Arbeit sind jedoch nur unmittelbare Klimaauswirkungen des Datenschutzes, d. h. solche Auswirkungen, die durch eine Veränderung des Verarbeitungssystems, insbesondere seiner Ausstattung und seiner physischen Parameter (Stromverbrauch) entstehen. Verhaltensänderungen und Fernwirkungen, z. B. insofern, als durch eine datenschutzrechtliche Unzulässigkeit Verarbeitungen nicht durchgeführt werden können, die ggf. auf die Erreichung von Klimaauswirkungen abzielen (vgl. dazu Coroama und Mattern [6]), werden in dieser Arbeit nicht betrachtet.

## 1.2 Szenario

Das in dieser Arbeit zu lösende Problem ist die Feststellung der Klimaauswirkungen des Datenschutzes. Da Klimaauswirkungen durch konkrete Emissionen konkreter Systeme entstehen, während der Datenschutz abstrakte Rechtssätze und Prinzipien aufstellt, bedarf es eines konkreten, wenn auch zulässigerweise hypothetischen Szenarios, das zwischen den beiden Bereichen vermittelt und insofern eine Ausgangslage für datenschutzrechtlich gebotene Veränderungen oder Einschränkungen und eine Grundlage für die Berechnung der Klimaauswirkungen schafft.

Für diese Zwecke ist es Absicht, dass das Szenario noch in verschiedener Hinsicht datenschutzrechtlich zu beanstanden ist und ggf. eher dem Wunschtraum einer Unternehmensleitung als der Datenschutzbeauftragten entspricht. Entsprechende datenschutzrechtlich gebotene Abschläge und Veränderungen erfolgen in späteren Teilen dieser Arbeit.

Dieser Arbeit liegt das folgende Szenario zugrunde:

Ein Unternehmen möchte eine Plattform anbieten, über die Patient*innen so einfach wie möglich Arzttermine buchen können, und zwar am Besten bei allen Praxen. Da das Unternehmen eine Ausgründung des HPI ist, soll die Plattform auf dem HPI-Rechenzentrum laufen. Patient*innen können auf der Plattform Termine bei kooperierenden, auf der Plattform eingelisteten Ärzt*innen buchen und gebuchte Termine ändern (Aufhebung, Verlegung). Sie werden an den Termin rechtzeitig erinnert und erhalten über die Plattform einen Buchungscode, mit dem sie sich bei der Anmeldung ausweisen können.

In einem nächsten Schritt möchte das Unternehmen die Plattform durch „KI“ anreichern. Damit sollen unnötige Arzttermine z. B. bei einer leichten Erkältung vermieden werden.

Hierfür soll vor die Terminbuchung ein Chatbot geschaltet werden, der den Patient*innen einige Fragen zu dem Grund ihres Termins bzw. ggf. zu ihren medizinischen Problemen stellt und dann entweder – in einfachen und unproblematischen Fällen wie der besagten leichten Erkältung – auf seine Diagnose und übliche alltägliche Behandlungsmittel hinweist, oder in allen anderen Fällen automatisch Ärzt*innen vorschlägt, bei denen man unmittelbar auf der Plattform einen Termin buchen sollte.

## 1.3 Plausibilität und Zweckmäßigkeit

Das konstruierte Szenario ist offensichtlich eine Hochrisikosituation für die Rechte und Freiheiten natürlicher Personen. In ihm fallen besonders schützenswerte Daten (medizinische Unterlagen) mit unvorhersehbaren technischen Systemen („KI“) zusammen.

Für jede der Kernfunktionen der Plattform aus dem Szenario gibt es bereits jetzt Systeme, die diese jeweilige Funktion bereitstellen. Soweit die Plattform die Buchung und Verwaltung von Terminen ermöglicht, entspricht dies im Wesentlichen den Funktionen der Anwendung „Doctolib“, die zudem wegen des Vorwurfs von Datenschutzproblemen in der Kritik steht.[4] Auch Selbstdiagnose-Tools z. B. in App-Form, die eine (vorläufige) Diagnose aufgrund von eigenen Angaben ermöglichen, gibt es bereits. Ein Beispiel ist die Gesundheitsanwendung „Ada“, welche, wie hier, über die Schnittstelle eines Chatbots mittels – laut eigenen Angaben – „KI“[5] funktioniert.

Daher ist es auch nicht unplausibel, dass ein Gesamtsystem errichtet wird, das all diese Aspekte vereint, zumal eine solche Systemkombination auch aus – rein – wirtschaftlicher Sicht zweckmäßig ist.

Das Szenario ist zudem auch insgesamt für diese Arbeit geeignet. Es enthält eine Vielzahl von Komponenten, bei denen Datenschutzmaßnahmen einschreiten müssen, wie dies z. B. offensichtlich bei dem Drängen zur Verwendung des Chatbots der Fall ist.

Zu berücksichtigen ist auch, dass das Szenario erst durch die große Anzahl von Datenschutzproblemen zweckmäßig wird, denn bei einem System, das nur im Hinblick auf einige wenige Punkte angepasst werden muss, kann eine verallgemeinerte Aussage schlechter getroffen werden, sodass das Endergebnis bzgl. der Emissionen mehr vom Zufall als vom Datenschutz abhängt.

## 1.4 Quantitative Annahmen

Damit die Emissionen des Systems sinnvoll sowohl vor als auch nach Anwendung der Datenschutzmaßnahmen bestimmt werden können, bietet es sich an, einige quantitative Annahmen zur Nutzung der Plattform zu treffen.

**Anzahl Ärzt*innen** Nach Angaben der Kassenärztlichen Bundesvereinigung gab es in Deutschland im Jahr 2021 $163,803$ ambulante Ärzt*innen [20].

[4] Vgl. `https://bigbrotherawards.de/2021/doctolib-gmbh`.
[5] Vgl. z. B. `https://ada.com/de/medical-quality/`.

**Anzahl und Art der Arzttermine** Die Kassenärztliche Bundesvereinigung berichtet weiterhin, dass die meisten Personen zwischen 3 und 5 Mal im Jahr einen Arzttermin wahrnehmen [18]. Wenn man dann bei einer Bevölkerung von $84432670$ Menschen [31] von je vier Terminen ausgeht, erhält man schätzungsweise $337730680$ jährliche, d. h. $938140$ tägliche Arztbesuche.

Eine Umfrage der Kassenärztlichen Bundesvereinigung hat zudem ergeben, dass rund $36\%$ aller Arztbesuche entweder aufgrund eines für sofort erhaltenen oder ohne Termin stattfinden können [19]. Im Umkehrschluss bedeutet dies, dass rund $64\%$ aller Arztbesuche aufgrund eines früher gebuchten Termins stattfinden, dies sind täglich $600409$ Arzttermine.

**Verbreitung** Unterstellt man, dass die Plattform (mittelfristig) von $5\%$ der Ärzt*innen genutzt wird, wird die Plattform nach vorgenanntem von $8190$ Ärztinnen bzw. Ärzten genutzt. Auf diese dürften dann rund $30020$ Terminbuchungen pro Tag entfallen.

Unterstellt man – deutlich überschätzend –, dass die durchschnittlich vier Arzttermine pro Person stochastisch unabhängig sind und einer zufälligen Auswahl aus allen Arztpraxen entsprechen, beträgt die Wahrscheinlichkeit, dass eine zufällige Person in einem Jahr kein Benutzerkonto auf der Plattform anlegt bzw. verwendet, $(95\%)^4 \approx 81.5\%$. Dies würde bedeuten, dass rund $18.5\%$ aller Personen innerhalb eines Jahres die Plattform genutzt haben werden, d. h. $15661732$ Personen in diesem Jahr ein Konto besitzen werden. Realistischer ist, dass jede Person, die ein Konto nutzt, pro Jahr 2 bis 3 Termine über die Plattform bucht. Damit erhält man einen plausibleren Wert von nur $4382920$ in einem Jahr aktiven Benutzerkonten.

**Terminsänderungen** Pauschalisiert wird für die Zwecke dieser Arbeit unterstellt, dass ca. $10\%$ der gebuchten Termine geändert (aufgehoben oder verlegt) werden.

**Nutzung und Effektivität des „KI“-Chatbos** Da nach dem Szenario der Chatbot „vor die Terminbuchung“ geschaltet werden soll, ist davon auszugehen, dass die Ablehnung der Nutzung nur umständlich möglich ist, sodass eine Vielzahl der Patient*innen den Chatbot nutzen werden. Daher ist anzunehmen, dass $90\%$ der Nutzer*innen den Chatbot verwenden.

Die Effektivität des Chatbots bestimmt sich zunächst nicht nach der Richtigkeit seiner Antworten. Da er nach dem Szenario unnötige Arzttermine vermeiden soll, bestimmt sich seine Effektivität danach, wie viele Terminbuchungen er vermeidet. Zu beachten ist, dass ein größerer Teil der Nutzenden von Anfang an fest plant, einen Termin zu buchen, z. B. aufgrund einer ärztlichen Überweisung, im Rahmen einer routinemäßigen Untersuchung oder aufgrund bekannter Symptome einer bekannten Erkrankung. In diesen Fällen wird der Chatbot von vornherein keine Terminbuchung verhindern können. Hinzu kommen von den übrigen Fällen alle, in denen keine milde und harmlose Erkrankung vorliegt bzw. ihr Vorliegen nicht ausgeschlossen werden kann, da in diesen Fällen der Chatbot eine Terminbuchung vorschlagen sollte. Daher erscheint es angebracht, anzunehmen, dass in höchstens $20\%$ (aller) Nutzungen der Chatbot verwendet wird und eine Terminbuchung ver-

meiden kann. In den restlichen 70% aller Fälle wird trotz seiner Benutzung dennoch ein Termin gebucht werden.

## 1.5 Nutzungsszenarien

Um Änderungen bei der Verwendung der Plattform, die sich aufgrund von Datenschutzmaßnahmen möglicherweise ergeben, untersuchen zu können, ist die Definition üblicher Nutzungsszenarien notwendig, die dann aufgrund der Maßnahmen modifiziert werden. Diese Nutzungsszenarien enthalten Angaben über fünf typisierte Ressourcen bzw. Aktionen: erforderliche HTTP-Anfragen, versendete E-Mails, Datenbank-Abfragen, Chatbot-Abfragen und Speicheraktionen. Diese Operationen werden dann bei den Klimaauswirkungen gesondert besprochen.

**I. Buchung eines Termins** Dies ist der standardmäßige und primäre Use Case des Szenarios: Nutzende haben ein medizinisches Anliegen oder Problem und möchten einen Termin buchen. Es wird angenommen, dass sie hierfür die folgenden Handlungen vornehmen:

1. Aufruf der Plattform (1 HTTP-Anfrage)
2. Anmeldung auf der Plattform (2 HTTP-Anfragen für Aufruf des Anmeldeformulars und dessen Absenden, 1 Datenbank-Abfrage für Authentifizierung)
3. Durchführung von ca. drei Suchanfragen (3 HTTP-Anfragen sowie 3 Datenbank-Abfragen)
4. Vorauswahl von ca. drei Ärzt*innen (3 HTTP-Anfragen sowie 3 Datenbank-Abfragen)
5. Suche nach Terminen bei der Vorauswahl (insgesamt 9 HTTP-Anfragen, je 3 je Auswahl; jeweils 9 DB queries)
6. Buchung eines Termins (3 HTTP-Anfragen für Termindaten, Kontrollseite und Buchungsbestätigung, 1 DB-Abfrage, sowie Speicherung eines Datensatzes)

Der Use Case umfasst damit insgesamt 21 HTTP-Anfragen, 17 Datenbank-Abfragen sowie die Speicherung eines Datensatzes im Umfang von ca. 306 Byte[6].

**II. Buchung eines Termins nach Verwendung des Chatbots** Dieser Use Case tritt nach den zuvor beschriebenen Annahmen in 70% der Fälle auf, nämlich dann, wenn nach Verwendung des Chatbots immer noch ein Termin zu buchen ist. Er hat den folgenden Ablauf:

1. die Schritte 1. und 2. des Use Cases I. (insg. 3 HTTP-Anfragen und 1 Datenbank-Abfrage)

[6] Unter der Annahme von je 8 Bytes für den Fremdschlüssel der Patientin oder des Patienten sowie den Fremdschlüssel der Arztpraxis, von je 17 Bytes für die Angabe des Terminsdatums und des Speicherdatums sowie von ca. 256 Bytes für weitere Informationen zum Termin; eine genauere Begründung derartiger Annahmen ist exemplarisch in Abschnitt 5.12. zu finden.

2. der Aufruf einer Informationsseite über den Chatbot (1 HTTP-Anfrage)
3. Kommunikation mit dem Chatbot bestehend aus ca. 26 Dialogbestandteilen (14 HTTP-Anfragen, davon 1 initiale und 13 für den Dialog; 13 Chatbot-Abfragen)
4. Anzeige eines Ergebnisses, d. h. die Empfehlung einer Suche (1 HTTP-Anfrage)
5. danach weiter mit 3. ff. des Use Cases I. (18 HTTP-Anfragen, 16 Datenbank-Abfragen sowie die Speicherung eines Datensatzes)

Der Use Case umfasst demnach insgesamt 37 HTTP-Anfragen, 17 Datenbank-Abfragen, 13 Chatbot-Abfragen und die Speicherung eines Datensatzes im Umfang von ca. 306 Bytes (s.o.).

**III. Abbruch nach Verwendung des Chatbots** Dieser Use Case tritt nach den zuvor beschriebenen Annahmen in $20\%$ der Fälle auf, nämlich dann, wenn nach Verwendung des Chatbots das Problem oder Anliegen gelöst ist, sodass ein Abbruch der Terminbuchung erfolgen kann. Er entspricht den Schritten 1. bis 4. des Use Cases II. und umfasst folglich insgesamt 19 HTTP-Anfragen, 1 Datenbank-Abfrage sowie 13 Chatbot-Abfragen.

**IV. Änderung eines Termins** Dieser Use Case umfasst die Verlegung sowie Absage (Änderung) eines bereits gebuchten Termins. Er läuft wie folgt ab:

1. die Schritte 1. und 2. des Use Cases Buchung eines Termins (insg. 3 HTTP-Anfragen und 1 Datenbank-Abfrage)
2. Aufruf einer Seite mit allen eigenen Terminen (1 HTTP-Anfrage und 1 Datenbank-Abfrage)
3. Aufruf der Seite für den zu ändernden Termin (1 HTTP-Anfrage und 1 Datenbank-Abfrage)
4. Auswahl der gewünschten Änderung und Eingabe zugehöriger Details wie z. B. den neuen Termin oder den Absagegrund sowie Bestätigung (1 HTTP-Anfrage)
5. Vornahme und Bestätigung des Eingangs der Terminsänderung (1 HTTP-Anfrage sowie 1 Datenbank-Abfrage)

Damit umfasst der Use Case 7 HTTP-Anfragen und 4 Datenbank-Abfragen.

**V. Erinnerung** Dieser Use Case umfasst die Versendung einer Erinnerungs-E-Mail wenige Tage vor dem Arzttermin. Dazu wird jeden Tag zu einer bestimmten Uhrzeit eine Datenbank-Abfrage ausgelöst, die alle für den Zieltag gebuchten Termine zurückgibt, für jeden Termin zwei weitere Abfragen ausführt, die die betroffene Person und die Arztpraxis zurückgeben, und anschließend eine E-Mail versendet.

**VI. Vor-Ort-Identifikation** Dieser Use Case umfasst die Identifikation vor Ort, d. h. in einer teilnehmenden Arztpraxis. Er läuft – eine vorherige Anmeldung unterstellt – wie folgt ab:

1. Anzeige einer Maske für die Eingabe des Buchungscodes (1 HTTP-Anfrage)

2. Suche nach dem eingegebenen Buchungscode und Anzeige des Termins, wenn vorhanden (1 HTTP-Anfrage, 1 Datenbank-Abfrage)
3. Bestätigung der Anwesenheit auf der angezeigten Seite (1 HTTP-Anfrage, 1 Datenbank-Abfrage)

Damit umfasst der Use Case 3 HTTP-Anfragen und 2 Datenbank-Abfragen.

## 1.6 Vorgehensweise

Die in dieser Arbeit behandelten Klimaauswirkungen des Datenschutzes sind die Auswirkungen auf das Klima, die aufgrund der Datenschutzmaßnahmen entstehen.

Durch diese Definition der Aufgabe kann das Problem auf zwei Schritte überführt werden: zunächst ist zu bestimmen, welche Datenschutzmaßnahmen bei dem Szenario anzuwenden sind, dann ist zu berechnen, welche Klimaauswirkungen die soeben bestimmten Datenschutzmaßnahmen mit sich bringen. Für jeden der beiden Schritte ist zunächst ein Maßstab aufzustellen und zu beschreiben, der anschließend auf das Szenario angewendet wird.

Diese Bachelorarbeit folgt dem so entstehenden Aufbau. In Kapitel 2 wird der Maßstab des Datenschutzes aufgestellt, in Kapitel 3 der Maßstab der Bemessung der Klimaauswirkungen. Dies wird anschließend auf das Szenario angewendet, und zwar in Kapitel 4 der Datenschutzmaßstab durch Analyse des Verarbeitungssystems und in Kapitel 5 durch Analyse der zuvor herausgearbeiteten Datenschutzmaßnahmen. Abschließend werden in Kapitel 6 die einzelnen Emissionsänderungen zusammengerechnet und bewertet.

# Kapitel 2

# Auswirkungen des Datenschutzes

Eine Untersuchung der „Klimaauswirkungen des Datenschutzes“ setzt schon begrifflich voraus, sich einen Überblick über die Auswirkungen des Datenschutzes insgesamt zu verschaffen, bevor dann geschaut werden kann, welche Klimafolgen einzelne Auswirkungen mit sich bringen.

Dafür werden in Kapitel 2.1 zunächst die Grundlagen des Datenschutzes skizziert, von denen anschließend nur die für diese Arbeit relevanten Grundsätze der Verarbeitung (in Kapitel 2.2) und die technisch-organisatorischen Maßnahmen (in Kapitel 2.3) genauer betrachtet werden sollen. Zudem wird in Kapitel 2.4 die Frage gestellt und beantwortet, ob und in welchem Umfang bestimmte Maßnahmen überhaupt Auswirkungen des Datenschutzes sind, und wie das jeweilige Fehlen von Zurechenbarkeit bei den Klimafolgen berücksichtigt werden kann.

## 2.1 Grundlagen des Datenschutzes

Datenschutz ist der Schutz vor der Verarbeitung personenbezogener Daten [32, Rn. 23ff.]. Er ist als Garant für die Entscheidungs- und Handlungsfreiheit, auch gerade in Bezug auf gegenwärtige oder zukünftige Minderheiten, in einer freiheitlich-demokratischen Gesellschaft unerlässlich [28]. Die konkrete und in der Praxis relevante Ausgestaltung erhält er durch das Datenschutzrecht, welches in der EU vorrangig durch die Datenschutz-Grundverordnung (DSGVO) sowie in deutlich geringerem Ausmaß und von untergeordneter Bedeutung durch nationale Rechtsvorschriften geregelt wird. Jene wurde von der Europäischen Union in dem Bestreben, ein einheitliches hohes Datenschutzniveau herzustellen (vgl. ErwGr 7 DSGVO), zugleich aber auch die Verwendung und den Austausch personenbezogener Daten im Rahmen dieses hohen Niveaus zu ermöglichen und zu fördern (vgl. ErwGr 2, 4, 6 und 7 DSGVO), im Jahre 2016 erlassen.

Nach der DSGVO bedarf jede Verwendung („Verarbeitung“) von Informationen, die sich auf identifizierte oder identifizierbare natürliche Personen beziehen („personenbezoge-

ne Daten“; vgl. Art. 4 DSGVO), einer Rechtsgrundlage, wie z. B. einem Vertrag, einer gesetzlichen Verpflichtung, einem besonderen berechtigten Interesse oder der Einwilligung der betroffenen Person (Art. 6 Abs. 1 DSGVO). Die Zwecke der Verarbeitung müssen vorab festgelegt werden; die Verarbeitung muss für diese Zwecke erforderlich sein und darf nicht für andere Zwecke erfolgen. Für besonders schützenswerte Daten wie die Daten von Kindern (Art. 8 DSGVO) oder bestimmte Kategorien besonders sensitiver Daten (Art. 9 und 10 DSGVO) gelten strengere Anforderungen.

Die betroffenen Personen sind über die Verarbeitung zu informieren (Art. 13 und 14 DSGVO). Sie haben das Recht, nach Maßgabe der Regelungen der DSGVO auf die Verarbeitung Einfluss zu nehmen: betroffene Personen erhalten auf Antrag Auskunft über die verarbeiteten personenbezogenen Daten; sie können die Berichtigung fehlerhafter Datensätze und in bestimmten Fällen auch die Beendigung oder Unterbrechung der Verarbeitung einschließlich ggf. der Löschung verlangen (Art. 15 bis 22 DSGVO).

An die Stellen, die personenbezogene Daten verarbeiten („Verantwortliche“), werden strenge Anforderungen gestellt. Beispielsweise haben sie bei der Verarbeitung auf die Sicherheit zu achten und Verletzungen zu melden (Art. 32 bis 34 DSGVO). Auch sind sie verpflichtet, sorgfältig darauf zu achten, mit wem sie zusammenarbeiten (Art. 26 und 28 DSGVO); in einigen Fällen müssen sie eine*n Datenschutzbeauftragte*n ernennen (Art. 37 bis 39 DSGVO) oder eine Abschätzung möglicher Folgen der Datenverarbeitung vornehmen (Art. 35 DSGVO).

Einschränkungen gelten auch, wenn personenbezogene Daten aus der Europäischen Union in Drittstaaten ausgeführt werden sollen, da dort möglicherweise kein ausreichender Schutz besteht (Art. 44 bis 49 DSGVO).

Zuletzt regelt die DSGVO noch Einzelheiten des Verfahrens und bestimmt Mittel zu ihrer Durchsetzung. Neben der Möglichkeit, vor den Gerichten Schadensersatz für datenschutzrechtswidriges Verhalten zu erstreiten (Art. 82 DSGVO), werden unabhängige Aufsichtsbehörden eingerichtet (vgl. Art. 58, 77, 83 und 84 DSGVO), die z. B. Verwarnungen und Anordnungen aussprechen sowie Bußgelder verhängen können (Art. 58 und 83 DSGVO).

Von diesen Vorschriften, die durch das Datenschutzrecht eingeführt werden, sind jedoch nicht alle für die Betrachtungen dieser Arbeit relevant. Auch wenn z. B. die Betroffenenrechte eine fundamentale Stütze des Datenschutzes sind [32, Rn. 72], werden diese in der Regel nur untergeordnete Auswirkungen auf die CO2-Emissionen haben, da für ihre Erfüllung weder zusätzliche Hardware anzuschaffen ist noch sich durch ihre Erfüllung der Rechenaufwand erheblich verändern dürfte. Das gleiche gilt für die Zuverlässigkeit von Auftragsverarbeitenden, die Benennung von Datenschutzbeauftragten, für die Vornahme einer Datenschutzfolgenabschätzung oder für die Verfahrens-, Sanktions- und Haftungsregelungen.

Soweit durch das Gebot der Verarbeitung in der EU die Verarbeitung nicht in einem ausländischen Rechenzentrum stattfinden kann, könnte zwar eine Auswirkung auf die CO2-Emissionen dadurch entstehen, dass das inländische Rechenzentrum weniger effizient oder aus einem anderen Grund klimaschädlicher ist. Es ist aber keine allgemeine Regel, dass inländische Rechenzentren klimaschädlicher als ausländische sind, weshalb es an der

Zurechenbarkeit zum Datenschutz fehlen dürfte (2.4), zudem kommt es auf diese Frage im Rahmen der Arbeit nicht an, da die Verarbeitung schon im Szenario in dem gut untersuchten HPI-Rechenzentrum stattfinden soll.

Daraus ergeben sich zwei Bereiche, die genauer zu betrachten sind: *die Grundsätze der Verarbeitung*, wie Rechtsgrundlage, Zweckbindung und Erforderlichkeitsgebot, (2.2) und *die technisch-organisatorischen Maßnahmen* zur operativen Sicherung des Datenschutzes (2.3).

## 2.2 Grundsätze der Verarbeitung

Nach Artikel 5 Absatz 1 und Artikel 6 DSGVO gelten die folgenden *verarbeitungsbeschränkenden Vorgaben*:

**Zweckbindung** Personenbezogene Daten müssen „für festgelegte, eindeutige und legitime Zwecke erhoben werden“ und sie „dürfen nicht in einer mit diesen Zwecken nicht zu vereinbarenden Weise weiterverarbeitet werden“ (Art. 5 Abs. 1 lit. b) DSGVO). Daraus ergibt sich einerseits, dass jede Verarbeitung personenbezogener Daten einen vorab bestimmten Zweck haben muss, sie also nicht „auf Vorrat “ erhoben werden dürfen, und die Daten grundsätzlich auch nur für diese Zwecke verwendet werden dürfen.

**Rechtsgrundlage** Zudem muss jeder Zweck, auf den eine Verarbeitung gestützt werden soll, durch eine „Rechtsgrundlage“ gerechtfertigt werden. Als Rechtsgrundlagen kommen nach Artikel 6 Absatz 1 Buchstaben a bis f DSGVO eine Einwilligung, die Erforderlichkeit für die Ausführung eines Vertrages, die Erforderlichkeit für die Erfüllung einer Rechtspflicht, die Erforderlichkeit für den Schutz lebenswichtiger Interessen, die Erforderlichkeit für die Wahrnehmung einer Aufgabe im öffentlichen Interesse sowie die Erforderlichkeit für die Wahrung überwiegender berechtigter Interessen in Betracht.

**Besonders schützenswerte Daten** Strengere Anforderungen gelten für besonders schützenswerte Daten, wie z. B. solche, aus denen „die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie [...] [genetischen] Daten, [biometrische] Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung“. Ihre Verarbeitung ist grundsätzlich verboten[7] und nur erlaubt, wenn – neben einer Rechtsgrundlage[8] – einer der Ausnahmetatbestände des Art. 9 Abs. 2 DSGVO anwendbar ist, wie bspw. eine ausdrückliche Einwilligung, die Erforderlichkeit für die Wahrnehmung arbeits- und sozialrechtlicher Rechte und Pflichten oder die offensichtliche Öffentlichmachung durch die betroffene Person. Zudem sind hier wegen des erhöhten Risikos strengere Anforderungen an die technisch-organisatorischen Maßnahmen (s. folgenden Unterabschnitt) zu stellen.

[7] So der ausdrückliche Wortlaut des Art. 9 Abs. 1 DSGVO.

[8] Strittig. Dafür z. B. Albers/Veit [1, Rn. 11], dagegen z. B. Mester [26]. Der Streitstand ist dargestellt bei Schulz [29, Rn. 5]. Hier wird die strengere Auslegung vertreten, u. a. damit nicht wegen Art. 9 Abs. 2 Buchst. e) tlw. ein geringeres Schutzniveau als für gewöhnliche Kategorien von personenbezogenen Daten besteht.

**Datenminimierung** Die Verarbeitung muss weiterhin „dem Zweck angemessen und erheblich sowie auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt sein“ (Art. 5 Abs. 1 lit. c) DSGVO). Dies bedeutet zum einen, dass als Rückversicherung für das Gebot der Zweckbindung die Verarbeitung für die Zwecke, auf die sie gestützt wird, sachdienlich sein muss (Zweckmäßigkeit), zum anderen dass sie für ihre Zwecke geeignet, erforderlich und angemessen sein muss (Verhältnismäßigkeitsprinzip). Eine Verarbeitung ist nicht (mehr) erforderlich, wenn es eine weniger eingreifende Verarbeitungsform gibt, die gleich geeignet ist. Sie ist unangemessen, wenn die Verarbeitung und ihre Zwecke auf der einen Seite sowie die Rechte und Interessen der betroffenen Personen auf der anderen Seite in einem Missverhältnis stehen.

**Speicherbegrenzung** Neben der sachlichen Beschränkung der Datenminimierung verlangt der Grundsatz der Speicherbegrenzung auch eine zeitliche Beschränkung. Personenbezogene Daten müssen „in einer Form gespeichert werden, die die Identifizierung der betroffenen Personen nur so lange ermöglicht, wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist“ (Art. 5 Abs. 1 lit. e) DSGVO). Unzulässig ist es daher, Daten auf Vorrat weiterzuspeichern, obwohl sie zurzeit nicht mehr berücksichtigt werden. Gleichwohl folgt hieraus nicht, die Daten unmittelbar nach Abschluss eines Vorgangs zu löschen, wenn berechtigte Gründe für eine weitere Speicherung vorliegen, wie z. B. gesetzliche Aufbewahrungspflichten oder die fachliche Notwendigkeit einer Protokollierung.

Daneben bestehen weitere zentrale Anforderung des Datenschutzrechts, nämlich dass die personenbezogenen Daten sachlich richtig zu sein haben, dass die Integrität und Vertraulichkeit der Daten gewahrt bleiben und dass die Verantwortlichen die Einhaltung ihrer Verpflichtungen nachzuweisen haben (vgl. Art. 5 Abs. 1 lit. d) und f) sowie Abs. 2 DSGVO). Diese – überwiegend nicht verarbeitungsbeschränkenden – Anforderungen gehören sachlich eher zu dem Kernbereich der technisch-organisatorischen Maßnahmen und nicht zur Rechtmäßigkeit an sich, sodass sie im folgenden Abschnitt als in den Gewährleistungszielen Integrität, Vertraulichkeit und Transparenz enthalten anzusehen sind.

## 2.3 Technisch-organisatorische Maßnahmen

Die DSGVO verlangt die Sicherung des Datenschutzes durch geeignete technisch-organisatorische Maßnahmen (Art. 24 DSGVO), die von den Verantwortlichen unter Beachtung des Stand der Technik, des Risikos für die Betroffenen und der Implementierungskosten ausgewählt werden.

Ein in der Praxis weit verbreiteter Ansatz zur Auswahl solcher Schutzmaßnahmen ist das *Standarddatenschutzmodell* [7] der Datenschutzkonferenz (gemeinsames Gremium der Landes- und Bundesdatenschutzbehörden). Das SDM vermittelt zwischen den rechtlichen Anforderungen und den technischen Implementierungen durch Einführung von sieben Gewährleistungszielen: Datenminimierung, Verfügbarkeit, Integrität, Vertraulichkeit, Nichtverkettung, Transparenz und Intervenierbarkeit.

Da die DSGVO eine gesetzlich verbindliche Regelung ist und die Gewährleistungsziele bloß eine Projektion dieser verbindlichen Regelungen darstellen, muss jedes dieser Ziele jeder-

zeit gewährleistet werden. Über die grundsätzliche Anforderung der Rechtmäßigkeit (s. o.) hinaus, welche sich auch in den Gewährleistungszielen wiederfindet, stellt das SDM generische Bausteine mit technisch-organisatorischen Maßnahmen bereit, welche diese Ziele unterstützen. Diese Hauptmaßnahmen sind[9]

- für das Ziel **Verfügbarkeit**

  die Redundanz von Daten, Systemen und Prozessen;

- für das Ziel **Integrität**

  Zertifikate; Tests gegen Soll-Werte; ein Datenschutz-Managementsystem; ein Rechtesystem;

- für das Ziel **Vertraulichkeit**

  die Verschlüsselung von Daten und Kommunikationsverbindungen; Authentifizierungsverfahren; ein Rechtesystem;

- für das Ziel **Transparenz**

  Spezifikation; Dokumentation; Protokollierung;

- für das Ziel **Nichtverkettung**

  Trennung; Zugriffsregelung; Anonymisierung und Pseudonymisierung;

- für das Ziel **Intervenierbarkeit**

  Unterstützung für das Löschen/Aufbewahren/Korrigieren/Einschränken von Daten; sowie Möglichkeit der Änderung von Verarbeitungen;

- und für das Ziel **Datenminimierung**

  Anwendung des „need-to-know"-Prinzips anstelle des "nice-to-have„-Prinzips; datenschutzfreundliche Voreinstellungen.

Bei diesen Hauptmaßnahmen handelt es sich nur um die üblichsten Maßnahmen, die dem jeweiligen Ziel im Interesse der Betroffenen Geltung verschaffen können. Weder muss jede einzelne Maßnahme in jedem System angewendet werden, noch garantiert die Anwendung aller dieser Maßnahmen pauschal Datenschutzkonformität [28].

Entscheidender Maßstab für die Auswahl der technisch-organisatorischen Maßnahmen ist das Risiko für die Rechte und Freiheiten der betroffenen Personen. Die DSGVO unterscheidet insoweit nur zwischen einem normalen und einem hohen Risiko; das Vorliegen keines Risikos ist bei der Verarbeitung von personenbezogenen Daten ausgeschlossen [28]. Entsprechend einer Ausarbeitung der Art. 29-Gruppe (Vorgängerinstitution des EDSA) haben sich 9 zentrale Indikatoren für ein hohes Risiko etabliert [3]: (1.) Bewerten oder Einstufen (Scoring), (2.) automatisierte Entscheidungsfindung mit Rechtswirkung oder ähnlich bedeutsamer Wirkung, (3.) systematische Überwachung, (4.) vertrauliche oder höchst persönliche Daten, (5.) Datenverarbeitung in großem Umfang, (6.) Abgleichen oder Zusammenführen von Datensätzen, (7.) Daten zu schutzbedürftigen Betroffenen, (8.) Innovative Nutzung oder Anwendung neuer technologischer oder organisatorischer Lösungen,

[9] Einteilung nach Rost [28, S. 126], kontrolliert und modifiziert anhand des SDM-Methodenhandbuchs der DSK [7, S. 30 ff.].

(9.) Betroffene werden an der Ausübung eines Rechts oder der Nutzung einer Dienstleistung bzw. Durchführung eines Vertrags gehindert. Liegen zwei oder mehr dieser Indikatoren vor, so muss ein hohes Risiko vermutet werden, jedoch kann in Einzelfällen auch eines dieser Indikatoren oder sonstige Umstände ausreichen, um ein hohes Risiko zu begründen.

Liegt ein hohes Risiko vor, dann müssen die Gewährleistungsziele stärker geschützt, d. h. die technisch-organisatorischen Maßnahmen konsequenter und umfassender umgesetzt werden [28].

## 2.4 Zurechenbarkeit

Für die Untersuchung, welche CO2-Emissionen *aufgrund des Datenschutzes* anfallen, stellt sich die Frage, ob eine bestimmte Maßnahme überhaupt ausschließlich *aufgrund* des Datenschutzes erforderlich wird, oder ob durchschnittliche vernünftige Verantwortliche eine solche Maßnahme auch ohne eine gesetzliche Pflicht umsetzen würden. Zudem stellt sich die Frage, wie das unterschiedliche Maß an Zurechenbarkeit in einer Gesamtrechnung bilanziert werden soll. Würde nämlich bei dieser nur auf solche Maßnahmen abgestellt, die ausschließlich aufgrund konkreter Vorschriften des Datenschutzrechts implementiert werden, so würde das Ergebnis die tatsächlichen Emissionen im Durchschnitt zu niedrig einschätzen, da nicht berücksichtigt würde, in welchen Teilen das konkrete Szenario ein Datenschutzniveau aufweist, das unter dem Stand der Technik liegt. Würden jedoch solche Emissionen voll berücksichtigt werden, so müssten viele Selbstverständlichkeiten wie bspw. die Bedachung des Rechenzentrums dem Datenschutz zugeschlagen werden, sodass das Ergebnis unberechtigt hoch wäre.

Da es sich bei der Nichtumsetzung des Datenschutzes um eine hypothetische und nicht umfassend untersuchbare Alternative zum Szenario handelt und deswegen generell keine unvoreingenommene Betrachtung erfolgen kann, wird eine sinnvolle Auflösung des Problems der Zurechenbarkeit nur durch eine typisierende Betrachtung erfolgen. Je wahrscheinlicher es ist, dass eine bestimmte Maßnahme auch ohne gesetzliche Pflicht implementiert werden würde, zum Beispiel weil sie zum Stand der Technik gehört oder offensichtlich im Interesse der Verantwortlichen ist, umso geringer hat der Anteil zu sein, zu dem die Emissionen der Maßnahme dem Datenschutz zugeschlagen wird.

Im Rahmen einer solchen typisierenden Betrachtung ist es wenig hilfreich, über wenige Prozentpunkte zu streiten, zumal eine definitive Festlegung in der Regel nicht möglich sein wird. Stattdessen bietet es sich an, eine Klasse von Kategorien zu errichten, denen eine Maßnahme anhand abstrakter Kriterien unter Angabe von Gründen zugewiesen werden kann.

Im Rahmen dieser Arbeit werden die folgenden vier Kategorien verwendet:

| Kategorie | Bilanzierung | Kriterien |
|---|---|---|
| **Ohne Zurechnung** | zu $0/3$ | für Maßnahmen, die von durchschnittlichen, verständigen Verantwortlichen offensichtlich auch ohne gesetzliche Pflicht für notwendig gehalten werden dürften |
| **Geringe Zurechnung** | zu $1/3$ | für Maßnahmen, die von durchschnittlichen, verständigen Verantwortlichen in der Regel auch ohne gesetzliche Pflicht für notwendig oder sehr vorteilhaft gehalten werden dürften |
| **Hohe Zurechnung** | zu $2/3$ | für Maßnahmen, die von durchschnittlichen, verständigen Verantwortlichen gegebenenfalls auch ohne gesetzliche Pflicht für notwendig oder vorteilhaft gehalten werden dürften |
| **Volle Zurechnung** | zu $3/3$ | für Maßnahmen, die von durchschnittlichen, verständigen Verantwortlichen nur aufgrund einer gesetzlichen Pflicht für notwendig oder vorteilhaft gehalten werden dürften |

Maßstab der jeweiligen Kategorien sind danach immer zwei Kriterien, einerseits, für wie essentiell eine Maßnahme auch ohne gesetzliche Pflicht gehalten wird (von *notwendig oder vorteilhaft* bis hin zu nur *notwendig*), sowie andererseits wie wahrscheinlich eine solche Einstufung ist (von *gegebenenfalls* bis hin zu *offensichtlich*).

Entscheidendes Indiz für das Ausmaß der Zurechnung kann in vielen Fällen auch sein, wie sehr die Maßnahme (auch) im Interesse des Verantwortlichen ist bzw. wie sehr sie zu diesem im Widerspruch steht. Je stärker ein solcher Widerspruch ausfällt, destso wahrscheinlicher ist es nämlich, dass die Maßnahme nur aufgrund des Datenschutzrechts vorgenommen wird.

# Kapitel 3

# Klimaauswirkungen

Bei der Untersuchung der CO2-Emissionen von Rechenzentren wird üblicherweise in die Bereiche „Embodied Carbon“ und „Operational Carbon“ unterteilt. Zu den Emissionen des ersten Bereichs zählt man solche, die auf den Auf- und Abbau des Rechenzentrums entfallen (wie z. B. Emissionen aufgrund von Materialverwendung oder Transport). Die Emissionen, die durch den Betrieb des Rechenzentrums, insbesondere den Stromverbrauch, entstehen, fallen dann in den zweiten Bereich.

Da der Datenschutz im Rahmen des vorliegenden Szenarios das Rechenzentrum in beiden Bereichen modifizieren kann, beispielsweise im ersten Bereich durch physische Schutzbarrieren und im zweiten Bereich durch Veränderungen in den ausgeführten Programmen oder den gespeicherten Daten, bietet es sich an, eine solche Unterteilung auch den Berechnungen dieser Arbeit zugrundezulegen.

Im Folgenden sollen die Berechnungsgrundlagen und -maßstäbe für die Klimaauswirkungen dargelegt und hergeleitet werden, bevor später das Szenario unter diese in Verbindung mit den zuvor dargestellten Datenschutzmaßstäben subsumiert werden kann.

## 3.1 Operational Carbon

Unter den datenschutzrechtlichen Begriff der Verarbeitung fallen die Erhebung, Speicherung, jegliche Verwendung sowie die Löschung von Daten. Alle diese Tätigkeiten fallen in den Bereich des Operational Carbon, da sie unmittelbar im Rahmen des laufenden Betriebs des Rechenzentrums stattfinden.

Der für diese Verarbeitungstätigkeiten erforderliche Aufwand lässt sich in zwei Gruppen einteilen: einerseits der Aufwand, der für die Vornahme von Berechnungen, d. h. die Ausführung von Computerprogrammen notwendig ist, andererseits der Aufwand für die Speicherung von Daten.

Da unterschiedliche Datenschutzmaßnahmen unterschiedlichen Einfluss auf beide Gruppen haben können (so liegt es auf den ersten Blick nahe, dass die technisch-organisatorische Maßnahme der Protokollierung mehr auf die zweite Gruppe Einfluss nimmt, als der

Grundsatz der Zweckbindung), sollen im Folgenden für beide Gruppen getrennt die CO2-Emissionsfaktoren bestimmt werden.

### 3.1.1 Energieverbrauch

Die Emissionen des Berechnungsaufwands und des Speicheraufwands entstehen – vorrangig – durch den Gebrauch von Energie für den Betrieb des jeweiligen Geräts. Wie viel CO2(-Äquivalent) durch den Verbrauch einer bestimmten Menge von Energie konkret emittiert wird, kann nicht pauschal gesagt werden, da dies von der Zusammensetzung des Stroms abhängt [10]. Wird mehr Strom aus erneuerbaren Energien wie Wind- oder Sonnenenergie erzeugt, so wird durch einen gleichartigen Betrieb des Geräts weniger CO2 emittiert, als wenn klimaschädliche Energiequellen wie Kohle oder Gas überwiegen [10].

Diese Zusammensetzung des Stroms ändert sich u. a. im Verlauf des Tages, da z. B. Sonnenenergie ganz überwiegend tagsüber produziert wird [10]. Damit ändert sich auch die Menge an Emissionen, die im Verhältnis zur verbrauchten Energie anfallen [10].

Für die Zwecke dieser Bachelorarbeit bietet es sich daher an, einen durchschnittlichen Emissionskoeffizienten zu verwenden. Betrachtet man den durchschnittlichen CO2-Koeffizenten für das HPI-Rechenzentrum über den gesamten Zeitraum seiner Berechnung, so erhält man ein Ergebnis von $0.402 \frac{\text{kg}}{\text{kWh}}$ [9].

### 3.1.2 Berechnungsaufwand

Die Ausführung von Computerprogrammen erfordert Energie. Dieser Berechnungsaufwand hängt im Wesentlichen von der Laufzeit des Programms und dem Energieverbrauch des Servers, auf dem das Programm ausgeführt wird, ab.

Für die Ermittlung des Berechnungsaufwands können zwei verschiedene Ansätze herangezogen werden. Einerseits könnte man den gesamten Zeitraum, in dem das Programm läuft, mit dem durchschnittlichen Energieverbrauch des Servers multiplizieren, andererseits könnte man die Zeit, in der das Programm tatsächlich im User- oder Systemmodus ausgeführt wird, mit dem Energieverbrauch des Servers auf Last verrechnen. Der zweite Ansatz erscheint hier vorzugswürdiger, denn Wartezeiten, die z. B. durch andere Prozesse oder Interrupts entstehen, können dem Programm in der Regel nicht zugerechnet werden und stehen zudem anderen Prozessen zur Benutzung offen. Außerdem ermöglicht dieser Ansatz einfacher die Berechnung von Zu- oder Abschlägen zu der Laufzeit sowohl in absoluter als auch prozentualer Hinsicht, da – potentiell variable – Wartezeiten nicht einberechnet werden müssen. Die Ermittlung der tatsächlichen Laufzeiten kann z. B. mit dem POSIX-Befehl `time` erfolgen.

Der Energieverbrauch des Servers hängt von einer Vielzahl von Faktoren, insbesondere von seiner Ausstattung ab. Da aus dem Szenario bekannt ist, dass die Plattform in dem HPI-Rechenzentrum laufen soll, bietet es sich an, die in dem HPI-Rechenzentrum am häufigsten verwendeten Servermodelle zu betrachten, da die Wahrscheinlichkeit einer Ausführung auf diesen am Höchsten ist. Eine Begehung des Serverraums im Rahmen des Bachelorprojekts hat drei identifizierbare Modelle ergeben, die mehr als zehn Mal verwendet

werden. Dies sind die Modelle *Fujitsu RX2530 M5*, *HPE XL225n Gen 10* und *DELL PowerEdge R430*.

Für die Modelle von Fujitsu und Dell finden sich Angaben zu ihrem Energieverbrauch. Nach den Angaben seines Datenblatts hat der Server von Fujitsu eine maximale Leistung von $883\mathrm{W}$ [12]. Das Produktkonfigurationstool von Fujitsu[10] gibt zudem an, dass das Servermodell alleine ohne jegliche Zusatzfunktionen $92\mathrm{W}$ leistet. Dell gibt für sein Servermodell an, dass es jährlich durchschnittlich $1760.3\mathrm{kWh}$ an Energie verbraucht [8], was einer Leistung von ca. $204\mathrm{W}$ entspricht. Der Blog IT Connected berichtet zudem, dass das Modell R430 im Leerlauf eine Leistung von $162\mathrm{W}$ und unter Last von $364\mathrm{W}$ erbringt [5].

Bei der Betrachtung dieser Werte ist zu beachten, dass die im Datenblatt für den Fujitsu-Server angegebene Leistung nur ein Maximalwert ist, der – wie man anhand des Konfigurationstools sieht – nur bei Ergänzung vieler Zusatzbestandteile erreicht wird. Zugleich fehlen dem Minimalwert viele notwendige Funktionen. Daher sind die beiden Werte als zu hoch bzw. zu niedrig anzusehen. Die Angaben zu dem Dell-Server erscheinen plausibel, auch im Vergleich mit anderen Server-Modellen. So hat bspw. das Modell *ProLiant DL380 Gen10* von HPE im durchschnittlichen Betrieb einen jährlichen Energieverbrauch von $1605\mathrm{kWh}$ [16], was in der Größenordnung des Dell-Modells liegt.

Da, wie zuvor begründet, nur der tatsächliche Berechnungsaufwand zu betrachten ist, d. h. die Leistung unter Last relevant ist, bietet es sich an, die Leistung unter Last aufgrund der obigen Werte auf einen Wert von $300\mathrm{W}$ zu schätzen. Mit dem zuvor berechneten CO2-Koeffizienten ergibt dies dann eine Emission von $120.6\mathrm{g}$ CO2-Äquivalent pro Stunde bzw. $33.5\mathrm{mg}$ pro Sekunde Berechnung.

### 3.1.3 Speicheraufwand

Die Emissionen für den Aufwand für die Speicherung von Daten lassen sich aus drei Gründen schwerer bestimmen als für den Berechnungsaufwand: Erstens ist der Speicher häufig in einem Server eingebaut, sodass nur Angaben für den gesamten Stromverbrauch des Servers, nicht aber gesondert für den Speicher bestehen. Zweitens geht es in vielen Forschungsarbeiten, die den Energieverbrauch von Speichern betreffen, um die Optimierung von Speicherverfahren und nicht um den eigentlichen Speicheraufwand. Drittens wird der Speicheraufwand häufig in Abhängigkeit von der Zeit angegeben und nicht in Abhängigkeit von der Datenmenge, was für die Zwecke dieser Arbeit jedoch sinnvoller wäre.

Zu beachten ist jedoch, dass die zuvor ermittelte Leistung des Servers auch den Zugriff auf den Speicher abdeckt. Unter der gleichen Annahme, dass nur die Zeit berechnet wird, die tatsächlich notwendig ist, kann daher der Speicheraufwand – zumindest für die Zwecke dieser Arbeit – abgeschätzt werden, wenn die Übertragungsrate für die Daten bekannt ist. Dann ergäben sich die Emissionen des Speicheraufwands im Verhältnis zur Datenmenge daraus, dass der Server so lange läuft, wie es für die Datenmenge aufgrund der Übertragungsrate erforderlich ist.

Vergleicht man die Datenblätter für die vorgenannten Server-Modelle, kann man erkennen, dass als Festplattenspeicher üblicherweise SSD-Festplatten über SATA bzw. SAS-An-

[10] Zu finden unter `www.fujitsu.com/configurator/public`.

schlüsse in einer Größe von 2,5 bzw. 3,5 Zoll erwartet werden. Sucht man nach Festplatten der jeweiligen Anbieter mit diesen Maßgaben, dann findet man zwei Gruppen von maximalen Übertragungsgeschwindigkeiten, einmal 0,6 GB/s (SATA) und einmal 1,2 GB/s (SAS).

Eine Auswahl von so gefunden Festplatten sind z. B.:

- HP Enterprise 872344-B21 – $0.6\mathrm{GB/s}$ (SATA)[11]
- HP Enterprise 875470-B21 – $0.6\mathrm{GB/s}$ (SATA)[12]
- HP Enterprise 875474-B21 – $0.6\mathrm{GB/s}$ (SATA)[13]
- Fujitsu S26361-F5782-L480 – $0.6\mathrm{GB/s}$ (SATA)[14]
- Fujitsu S26361-F5787-L480 – $0.6\mathrm{GB/s}$ (SATA)[15]
- HP Enterprise 762751-001 – $1.2\mathrm{GB/s}$ (SAS)[16]
- HP Enterprise P9M80A – $1.2\mathrm{GB/s}$ (SAS)[17]
- HP Enterprise N9Z14A – $1.2\mathrm{GB/s}$ (SAS)[18]

Da die Produktangaben nur einen Höchstwert darstellen und da der Wert $0.6\mathrm{GB/s}$ in der Recherche etwas häufiger auftauchte, als der Wert $1.2\mathrm{GB/s}$, wird im Rahmen dieser Arbeit angenommen, dass der Zugriff auf Speichermedien mit einer Geschwindigkeit von $0.6\mathrm{GB/s}$ erfolgt. Wenn man dies mit dem im vorherigen Abschnitt hergeleiteten Faktor der CO2-Emission in Abhängigkeit zur Zeit ins Verhältnis setzt, erhält man einen CO2-Verbrauch für den Speicheraufwand von $\frac{33.5\mathrm{mg/s}}{0.6GB/\mathrm{s}}$, d. h. $55.83\mathrm{mg/GB}$.

### 3.1.4 Typisierte Operationen

Bei einer Web-Plattform, wie sie in dem Szenario vorliegt, fallen bestimmte Operationen regelmäßig an und ermöglichen so eine Abstraktion bei der Beschreibung der Systemtätigkeit, die Veränderungen in dem System geeignet und einfach darstellen lässt. Daher bietet es sich an, einige typisierte Operationen zu definieren und für diese die CO2-Emissionen zu bestimmen.

**HTTP-Anfragen** Die fundamentale Operation bei einer Web-Plattform ist die HTTP-Anfrage. Diese besteht daraus, dass der Webbrowser eine Anfrage, ggf. mit weiteren Daten

[11] Vgl. `https://www.jacob.de/produkte/hpe-480gb-sata-872344-b21-artnr-3411947.html`, Stand: 22.09.2023.
[12] Vgl. `https://www.jacob.de/produkte/hpe-mixed-use-875470-b21-artnr-3749087.html`, Stand: 22.09.2023.
[13] Vgl. `https://www.jacob.de/produkte/hpe-mixed-use-875474-b21-artnr-3743598.html`, Stand: 22.09.2023.
[14] Vgl. `https://www.jacob.de/produkte/fujitsu-ssd-sata-s26361-f5782-l480-artnr-6794325.html`, Stand: 22.09.2023.
[15] Vgl. `https://www.jacob.de/produkte/fujitsu-ssd-m-2-s26361-f5787-l480-artnr-6794304.html`, Stand: 22.09.2023.
[16] Vgl. `https://www.jacob.de/produkte/hewlett-packard-sps-drv-762751-001-artnr-2331414.html`, Stand: 22.09.2023.
[17] Vgl. `https://www.jacob.de/produkte/HEWLETT-PACKARD-ENTERPRISE-P9M80A-artnr-3013891.html`, Stand: 22.09.2023.
[18] Vgl. `https://www.jacob.de/produkte/HEWLETT-PACKARD-ENTERPRISE-N9Z14A-artnr-3017216.html`, Stand: 22.09.2023.

z. B. aus einem Formular, an den Server sendet und der Server dann die Anfrage mit einer weiteren Webseite beantwortet. In Folge der Antwort werden dann auch weitere – statische – Ressourcen wie CSS-Dateien, JavaScript-Programme oder Bilder übertragen.

Für eine HTTP-Anfrage fällt überwiegend nur Berechnungsaufwand an. Die Bearbeitungsdauer einer Anfrage hängt von einer Vielzahl von Faktoren ab, wie der gewählten Programmierumgebung und der Komplexität der Webseitengenerierung. Dennoch ist eine – grobe – Schätzung anhand mehrerer Beispiele möglich: Zunächst wurden die `development.log`-Dateien von drei lokal auf meinem Computer ausgeführten Ruby on Rails-Anwendungen (ein kleines privates Projekt, die QPixel-Software[19] sowie die Xikolo-Anwendung von openHPI) betrachtet. Diese Log-Dateien enthalten Informationen über die Gesamtzeit und die für Datenbankabfragen verwendete Zeit, sodass diese einfach rausgerechnet werden können. Es ergeben sich dabei Medianzeiten von $19.1\text{ms}$, $35.2\text{ms}$ und $570.3\text{ms}$. Zudem wurden Anfragen an die Webseiten von Google, Doctolib und Spiegel gesendet, bei denen, die Round-Trip-Time nach der Ausgabe des `ping`-Befehls abgezogen, Gesamtzeiten von $273\text{ms}$, $200\text{ms}$ bzw. $257\text{ms}$ für die Anfrage nur der Webseite herauskamen. Hier ist jedoch noch der Aufwand für Datenbankabfragen enthalten.

Hinzu kommt ein Berechnungsaufwand für durchschnittlich mindestens vier oder fünf statische Ressourcen. Diese kann der Server dank Caching deutlich schneller bereitstellen. Nimmt man aufgrund der vorstehenden Ergebnisse durchschnittlich ca. $100\text{ms}$ für die Erzeugung der Webseite (ohne Datenbank) und für jede der Ressourcen ca. $20\text{ms}$ an, ergibt dies einen geschätzten Berechnungsaufwand von $200\text{ms}$ für jede HTTP-Anfrage.

Diesem Berechnungsaufwand entspricht nach den vorstehenden Maßstäben einer Emission von ca. $6.7\text{mg}$ CO2-Äquivalent.

**E-Mail-Versand** Für die Zwecke dieser Arbeit wird angenommen, dass der Aufwand für das Versenden einer E-Mail dem Aufwand einer HTTP-Anfrage entspricht. Dieser Annahme liegt die Beobachtung zugrunde, dass E-Mails, insbesondere im kommerziellen Kontext, heutzutage häufig ebenfalls mit HTML, CSS und Bildern formattiert sind und somit als „kleine Webseiten“ angesehen werden können. Der Aufwand für die Erstellung und die Versendung aus dem lokalen Netzwerk dürfte daher als ähnlich angesehen werden.

**Datenbank-Abfragen** Die Daten der Plattform, insbesondere Arztdaten, Kontodaten und Terminsdaten werden in einer Datenbank gespeichert. Abfragen werden erforderlich, wenn auf diese Daten lesend oder schreibend zugegriffen werden soll. Es fällt hierbei sowohl Berechnungs- als auch Speicheraufwand an. Hierzwischen kann jedoch nach außen nur schlecht unterschieden werden. Wegen der Art und Weise der Definition des Speicheraufwands als Spezialfall des Berechnungsaufwands kann jedoch die Gesamtzeit der Ausführung einer Datenbank-Abfrage für die Berechnung des Gesamtaufwands verwendet werden. Unter Auswertung der bereits zuvor genannten Log-Dateien ergeben sich bei insgesamt $214102$ protokollierten Abfragen Durchschnittszeiten von $0.76\text{ms}$ für das Erzeugen von Datensätzen, $1.09\text{ms}$ für das Ändern von Datensätzen, $0.62\text{ms}$ für das Laden von Datensätzen und $0.52\text{ms}$ für das Löschen von Datensätzen.

[19] vgl. `https://github.com/codidact/qpixel`

Aufgrund dieser Ergebnisse wird für die Zwecke dieser Arbeit die Ausführungszeit für eine Datenbankabfrage auf den mittleren Wert von $0.75\text{ms}$ geschätzt. Dieser Zeit entspricht eine CO2-Emission von ca. $25.125\text{mg}$ pro Abfrage.

**Chatbot-Abfragen** Die Kommunikation mit dem Chatbot wird über Chatbot-Abfragen modelliert. Diese funktionieren so, dass an den Chatbot ein bestimmter Text gesendet, der Text verarbeitet, und anschließend ein weiterer Text als Antwort zurückgegeben wird. Es fällt überwiegend nur Berechnungsaufwand an.

Der wesentliche Faktor für die Berechnungsdauer ist die Ausgabelänge [22]. Eine Benchmarking-Untersuchung ergibt Berechnungsdauern von ca. 2 bis ca. 15 Sekunden auf verschiedener Hardware und mit verschiedenen Aufgabenstellungen [36]. Dies zugrundegelegt wird geschätzt, dass eine Chatbot-Abfrage ca. 3 Sekunden für die Beantwortung einer Abfrage benötigt, denn die in der Benchmark verwendete Fallgruppe, dass aus einer potentiell mittellangen Eingabe eine mittlere bis mittelkurze Ausgabe erzeugt werden soll, erscheint für das hier verwendete System am plausibelsten, da der Chatbot Rückfragen stellen und dann eine kurze Antwort geben soll. Umfassende Texterzeugung ist nicht beabsichtigt. Diese Fallgruppe wies Bearbeitungszeiten zwischen 2 und 4 Sekunden auf verschiedenen Hardware-Systemen auf. Damit ist anzunehmen, dass eine einzelne Chatbot-Abfrage eine Emission von ca. $100.5\text{mg}$ CO2-Äquivalent erzeugt.

Eine Bilanzierung des Berechnungsaufwands für das Training des „KI“-Modells war ungeachtet der später noch zu besprechenden Punkte schon deshalb nicht veranlasst, weil ein einmal erzeugtes Modell gespeichert und beliebig vervielfältigt werden kann, weshalb eine Erhöhung der Anzahl der Chatbot-Abfragen nicht eine Erhöhung des Trainingsaufwands mit sich bringt.

### 3.1.5 Zusammenfassung

Aufgrund der vorstehenden Erwägungen und Untersuchungen werden für die Zwecke der Bestimmung des Operational Carbons in dieser Arbeit die folgenden Arten von Aufwand mit den zugehörigen Emissionsfaktoren verwendet:

| Aufwand | Einheit | CO2eq-Emissionen je Einheit |
|---|---|---|
| Rechenzeit | 1s | 33.5mg |
| Speicherzugriff | 1GB | 55.83mg |
| HTTP-Anfrage | 1 | 6.7mg |
| E-Mail-Versand | 1 | 6.7mg |
| Datenbank-Abfrage | 1 | 25.125mg |
| Chatbot-Abfrage | 1 | 100.5mg |

## 3.2 Embodied Carbon

In der Kategorie des Embodied Carbon werden Emissionen zusammengefasst, die außerhalb des operativen Betriebs eines Rechenzentrums entstehen, wie beispielsweise bei des-

sen Auf- und Abbau, für die Erzeugung oder Förderung verwendeter Materialien, den Transport oder das Recycling.

Das Rechenzentrum, in dem das in dieser Arbeit betrachtete Szenario läuft, existiert jedoch schon – sowohl in tatsächlicher Hinsicht als auch in dem Szenario. Da vorliegend nur Differenzen betrachtet werden, die zwischen dem im Szenario beschriebenen Ausgangssystem und dem Szenario in seiner Gestalt nach der Einwirkung durch die Datenschutzmaßnahmen bestehen, kann Embodied Carbon in diesem Szenario nur anfallen, wenn konkrete physische Veränderungen an dem Szenario erforderlich werden. Eine solche Veränderung könnte beispielsweise sein, dass um einen einzelnen Server eine besonders sichere Zugangsschleuse eingebaut werden muss.

In diesen Fällen kann der Anfall von Embodied Carbon mittels des in der Industrie anerkannten PAIA-Verfahrens [27, 24] geschätzt werden. Inzwischen stellen viele Anbieter im Enterprise IT-Bereich zudem auch so genannte *Life Cycle-Assessments* bereit, in der sie selber eine Schätzung vornehmen und die Ergebnisse darstellen. Ein solches Assessment wurde beispielsweise von Dell für den zuvor bereits beschriebenen *DELL PowerEdge R430* bereitgestellt [8].

# Kapitel 4

# Analyse des Verarbeitungssystems

Um ihre Klimaauswirkungen zu bestimmen, ist es zunächst erforderlich, festzustellen, welche Datenschutzmaßnahmen die zuvor dargestellten Grundsätze bei Anwendung auf das gegenständliche Verarbeitungssystem gebieten. Hierzu bietet sich das Instrument der „Datenschutz-Folgenabschätzung“ (kurz: DSFA) an, da ihre Vornahme weitestgehend der nach den vorgenannten Grundsätzen vorzunehmenden Analyse entspricht: Zunächst sind die Verarbeitungsvorgänge zu beschreiben, dann ist die Zulässigkeit der Verarbeitung zu prüfen, anschließend sind die Risiken zu identifizieren, und zuletzt sind Abhilfemaßnahmen für diese Risiken und die Bedenken bei der Zulässigkeit vorzuschlagen.

Für das zuvor beschriebene Verarbeitungssystem habe ich eine Datenschutz-Folgenabschätzung vorgenommen. Das vollständige Ergebnis ist im Anhang A zu finden, im Folgenden werden die wesentlichen Erkenntnisse dargestellt.

## 4.1 Verarbeitungsvorgänge

Aufgrund des obigen Szenarios können acht Verarbeitungsvorgänge identifiziert werden, nämlich die Vorgänge

1. Anzeige von Arztinformationen (Termin),
2. Buchung eines Termins (Termin),
3. Verwaltung eines Termins (Termin),
4. Terminerinnerung (Termin),
5. Vor-Ort-Identifizierung (Termin),
6. Erzeugung von Trainingsdaten (Chatbot),
7. Training (Chatbot) sowie
8. Konsultation (Chatbot).

Jedem dieser Vorgänge können ein oder mehrere Zweck zugewiesen werden, für welche sie durchgeführt werden. Der Zweck ist bei den Vorgängen 1 und 2 die Vermittlung von Patient*innen, bei Vorgang 2 zusätzlich die Geschäftsanbahnung, bei Vorgang 4 ihre Sicherung, bei Vorgang 3 die Ermöglichung von Flexibilität für die Patient*innen, bei Vorgang 5 die Wahrnehmung des Termins, bei den Vorgängen 6 und 7 die Ermöglichung/Verbesserung des Chatbots und bei Vorgang 8 die Vereinfachung für Patient*innen sowie die Entlastung des medizinischen Personals.

## 4.2 Zulässigkeit

Die Verarbeitungsvorgänge 1 bis 5 sowie der Verarbeitungsvorgang 8 können allesamt auf eine Rechtsgrundlage im Sinne des Art. 6 DSGVO sowie die Vorgänge 2 bis 5 und 8, die Gesundheitsdaten betreffen, auch auf einen Ausnahmegrund im Sinne des Art. 9 DSGVO gestützt werden. Hieraus ergeben sich jedoch teilweise organisatorische Auflagen, die der DSFA entnommen werden können.

Die Verarbeitungsvorgänge 6 und 7 können zwar auf eine Rechtsgrundlage im Sinne des Art. 6 DSGVO gestützt werden, nicht jedoch auf einen Ausnahmegrund im Sinne des Art. 9 DSGVO. Eine Einwilligung scheidet aus, da die Anforderung der Widerruflichkeit nicht umsetzbar ist. Es kann nicht mit der hinreichenden Sicherheit und mit zumutbarem Aufwand durchgesetzt werden, dass personenbezogene Daten nicht mehr Teil des KI-Modells sind. Die weiteren Ausnahmegründe scheitern daran, dass ihr Tatbestand offensichtlich nicht erfüllt ist, dass die Verarbeitung nicht im Rechtssinne erforderlich ist bzw. dass die Verarbeitung zu rein privat-kommerziellen Zwecken und nicht im öffentlichen Interesse erfolgt. Damit verstoßen diese Verarbeitungsvorgänge gegen das Verarbeitungsverbot des Art. 9 Abs.1 DSGVO; sie sind folglich **rechtswidrig** und dürfen nicht durchgeführt werden.

Im Rahmen dieser Bachelorarbeit wird davon abgesehen, als Folge dieser Rechtswidrigkeit auch den Verarbeitungsvorgang 8 zu streichen, da dieser mit einem datenschutzkonform erstellten Modell oder mit einer anderen Technik (z. B. Entscheidungsbäume) weiterhin sinngemäß möglich ist. Die Frage, ob und wie eine solche Lösung umgesetzt wird, ist eine betriebliche und keine des Datenschutzes, sodass ihre Beantwortung auch nicht dem Datenschutz zugerechnet werden kann.

Soweit die Verarbeitungsvorgänge auf eine Rechtsgrundlage gestützt werden können, entsprechen sie auch grundsätzlich den Anforderungen von Notwendigkeit und Angemessenheit. Diese Beurteilung könnte jedoch auch anders ausfallen, wenn die Marktmacht der Plattform zu einem faktischen Benutzungszwang führen würde.

## 4.3 Risiken

Für die zulässigen Verarbeitungsvorgänge können insgesamt 26 Risiken identifiziert werden (die Risiken 1.a. bis 8.g. in der DSFA). Ein Risiko äußert sich dabei so, dass unter Verstoß gegen eines oder mehrere Gewährleistungsziele ein Nachteil für die betroffenen Perso-

nen eintritt. Beispielsweise wurde für Verarbeitungsvorgang 2 das Risiko 2.a identifiziert, dass „unter Verletzung des Gewährleistungsziels der Vertraulichkeit und der Nichtverkettung hochsensitive Gesundheitsdaten Unbefugten bekannt und von diesen ggf. weiterverarbeitet werden könnten“; für Verarbeitungsvorgang 4 wurde z. B. das Risiko 4.a. identifiziert, dass „unter Verstoß gegen das Gewährleistungsziel der Integrität und der Verfügbarkeit eine Terminerinnerung mit falschen Angaben oder nicht versendet werden und daher im Vertrauen auf die Erinnerung bzw. das Senden der Erinnerung ein Termin nicht oder nicht richtig wahrgenommen werden könnte“; für Verarbeitungsvorgang 8 schließlich wurde u. a. das Risiko 8.g. identifiziert, dass es „unter Verstoß gegen die Intervenierbarkeit nicht möglich sein könnte, ohne Inanspruchnahme des Chatbots einen Termin zu buchen“.

Dabei verteilen sich die Risiken der zulässigen Verarbeitungsvorgänge (VV) 1 bis 5 sowie 8 auf die Gewährleistungsziele wie folgt:

| Ziel | VV 1 | VV 2 | VV 3 | VV 4 | VV 5 | VV 8 |
|---|---|---|---|---|---|---|
| Verfügbarkeit | | 2.b. | | 4.a. | 5.a., 5.c. | 8.d. |
| Integrität | 1.a., 1.b. | | 3.a | 4.a., 4.b. | 5.a. | 8.a. |
| Vertraulichkeit | | 2.a. | 3.b. | 4.b. | 5.b. | 8.b. |
| Transparenz | 1.b. | 2.d. | 3.c. | | 5.b. | 8.c. |
| Nichtverkettung | 1.c. | 2.a. | 3.b. | 4.b. | 5.b. | 8.b. |
| Intervenierbarkeit | 1.a. | 2.e. | | 4.d. | 5.c. | 8.f., 8.g. |
| Datenminimierung | 1.c., 1.d. | 2.c. | | 4.c. | | 8.e., 8.f. |

Sodann wurde für jeden Verarbeitungsvorgang unter Berücksichtigung der identifizierten Risiken eine Schwellwertanalyse nach den Kriterien von WP248 [3] durchgeführt. Hierbei ergab sich für die Verarbeitungsvorgänge 2, 3 und 8 ein **hohes Risiko**, während das Risiko der übrigen Vorgänge (noch) als normal eingestuft werden konnte.

## 4.4 Abhilfemaßnahmen

Für die Verarbeitungsvorgänge wurden dann in drei Schritten die Abhilfemaßnahmen vorgeschlagen. Zunächst wurden alle Auflagen, die sich aus der Beurteilung der Zulässigkeit ergeben, wie bspw. die Sicherstellung der Freiwilligkeit bestimmter Verarbeitungsvorgänge, als Abhilfemaßnahmen vorgeschlagen.

Sodann wurden zunächst für jedes Risiko die generischen technisch-organisatorischen Maßnahmen aus dem Standarddatenschutzmodell [7] für die jeweils verletzten oder gefährdeten Gewährleistungsziele betrachtet und diejenigen Maßnahmen ausgewählt, die nicht offensichtlich ungeeignet scheinen, dem jeweiligen Risiko zu begegnen. Als offensichtlich ungeeignet wurden hierbei insbesondere solche Maßnahmen angesehen, die sich ausschließlich auf eine andere Ausprägung des jeweiligen Gewährleistungsziels beziehen. Kommt es bspw. auf die Integrität von Prozessen an, so sind Software-Tests geeignet, jedoch das Löschen oder Berichtigen falscher Daten eher weniger, da dieses sich auf die Integrität von Daten bezieht.

Die so für jedes Risiko eines Verarbeitungsvorgangs ausgewählten generischen Maßnahmen werden dann zu spezifischen Maßnahmen für einen Verarbeitungsvorgang zusammengefasst. Damit wurde einerseits die Doppelung von Maßnahmen vermieden, andererseits wurden zusammengehörige Komponenten zusammengefügt. Beispielsweise wurden die Redundanz von Hard- und Software (4.a.3.), die Umsetzung von Reparaturstrategien und Ausweichprozessen (4.a.4.) und die Erstellung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit (4.a.5.) zu der spezifischen Maßnahme Herstellung von Redundanz von Hard- und Software sowie Infrastruktur sowie Festlegung und regelmäßige Kontrolle eines Verfahrens zur Wiederherstellung der Verarbeitungstätigkeit (4.B.) zusammengefasst, da die drei Verarbeitungsvorgänge im Wesentlichen den gleichen Gegenstand betrafen, nämlich die Wiederherstellbarkeit durch Redundanz.

Hiernach ergeben sich dann z. B. gegen das oben beschriebene Risiko 8.g. für die Intervenierbarkeit die folgenden beiden technisch-organisatorischen Maßnahmen:

- Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten und Schaffung notwendiger Datenfelder zur Umsetzung dieser Maßnahmen (8.N.)
- Gewährleistung der Möglichkeit, die Hauptanwendung zu nutzen, wenn der Chatbot nicht funktioniert, die Einwilligung verweigert oder widerrufen oder die Nutzung auf andere Weise abgelehnt wird (8.O.)

Eine vollständige Auflistung der vorgeschlagenen TOM ist auch in der Anlage 1 zur DSFA zu finden.

Für die Zwecke der folgenden Analyse der Abhilfemaßnahmen sind diese anhand ihres wesentlichen Charakters in dreizehn Kategorien gegliedert, nämlich in die Kategorien ausschließlich organisatorische Maßnahmen, unzulässige Verarbeitungen, physische Schutzmaßnahmen, die Abhärtung des Systems, Redundanz und Wiederherstellbarkeit, Gewährleistung durch Software/Konfiguration, Änderungen in der Benutzeroberfläche, regelmäßige Abfragen, Authentifizierung und Autorisierung, Verschlüsselung, Trennung, Protokollierung, sowie Software-Tests.

# Kapitel 5

# Analyse der Datenschutzmaßnahmen

Im Folgenden sollen die Maßnahmen, die nach den Feststellungen des vorherigen Kapitels aufgrund des Datenschutzes geboten sind, auf ihre Klimafolgen analyisert werden. Die verschiedenen spezifischen Einzelmaßnahmen werden nach ihrer Art in bestimmte Kategorien gruppiert und dann gruppenweise betrachtet. Die Bezeichnungen richten sich nach der bereits erwähnten DSFA.

Soweit Datenschutzmaßnahmen durch Anpassung der Use Cases modelliert werden können, kann ein Use Case $x$ wie folgt beeinflusst werden: sie können die Anzahl der Operationen $s_{x,o}$ (z. B. HTTP-Anfragen) um eine bestimmte Zahl $\Delta s_{x,o}$ erhöhen oder verringern (Beispiel: wenn ein Zwischenschritt eingeführt wird); sie können die Anzahl der Ausführungen $N_x$ des Use Cases um einen Faktor $\Delta N_x$ verändern. Zudem können Datenschutzmaßnahmen den Aufwand je Operation $A_e$ um einen bestimmten Betrag $\Delta A_e$ verändern. Danach ergibt sich der Aufwand für eine typisierte Operation bei einem Use Case nach Anwendung aller Datenschutzmaßnahmen über die folgende Formel:

$$E(x, o) = (s_{x,o} + \Delta s_{x,o}) * N_x * \Delta N_x * (E_o + \Delta E_o)$$

Hierbei sind die Variablen $s$ und $\Delta A_e$ über allen Maßnahmen zu addieren und die Variable $N$ zu multiplizieren.

Außerhalb der Modifikation von Use Cases kann eine Datenschutzmaßnahme z. B. durch die Ergänzung eines neuen Use Cases oder durch fixe bzw. zeitabhängige Veränderungen bei den CO2-Emissionen wirken.

## 5.1 Ausschließlich organisatorische Maßnahmen

Die Maßnahmen 1.C., 2.B., 2.I., 3.A.B., 3.B., 5.A., 5.F., 5.G., 8.D., 8.E., 8.J. und 8.M. sind ausschließlich organisatorische Maßnahmen. Dies bedeutet, dass sie keine technische Verän-

derung mit sich bringen; stattdessen handelt es sich um Maßnahmen wie Dokumentation oder die Anpassung von Betriebsprozessen. Ihnen lassen sich im Rahmen des in dieser Arbeit verfolgten Ansatzes (vgl. Kapitel 4) keine besonderen Emissionen zuordnen. Dies ist zudem die erwartete Folge, denn es besteht z. B. kein allgemeines Gesetz, dass eine Mitarbeiterin, deren Zuverlässigkeit festgestellt wurde und die förmlich auf Geheimhaltung verpflichtet ist, mehr oder weniger Emissionen verantwortet als ein Mitarbeiter, bei dem dies nicht geschehen ist.

## 5.2 Unzulässige Verarbeitungen

Die Verarbeitungsvorgänge 6 und 7, d.h. die Erzeugung von Trainingsdaten und das Training eines Chatbots sind datenschutzrechtlich unzulässig und daher insgesamt zu unterlassen. Dieses Ergebnis ist dem Datenschutz auch voll zuzurechnen, da es zu dem ursprünglichen Konzept des Verantwortlichen in eklatantem Widerspruch steht, sodass nicht davon auszugehen ist, dass Verantwortliche diese Maßnahme freiwillig ergriffen hätten.

Wesentlicher Faktor bei dem Training einer „KI“ ist der Rechenaufwand. Das Szenario enthält keine spezifischen Aussagen, wie das Modell funktionieren soll. Daher bietet es sich an, den Vergleich mit anderen „KI“-Modellen, insbesondere für Chatbots, zu suchen[20]. Facebook hat für die Entwicklung seiner Modelle LLaMa und LLaMa 2 Paper veröffentlicht. Laut diesen Angaben seien für das Training von LLaMa (in allen Varianten) ca. $1015\text{t}$ CO2-Äquivalent [35] und für das Training von LLaMa 2 ca. insgesamt $539\text{t}$ CO2-Äquivalent [34] angefallen. Für LLaMa 2 sind zudem Emissionen in Abhängigkeit von der Größe der verschiedenen Modellvarianten angegeben. Für das kleinste Modell mit 7 Mrd. Parametern seien danach $31.22\text{t}$ emittiert worden, für das größte Modell mit 70 Mrd. Parametern $291.42\text{t}$. Mangels spezifischerer Angaben im Szenario lässt sich annehmen, dass ein Modell mittlerer Art verwendet wird, sodass Emissionen zwischen $60\text{t}$ und $150\text{t}$ anfallen. Vergleichbare Werte liefert auch die Holistic Evaluation of Language Models von Liang et al. [22].

Dies zugrundegelegt kann man annehmen, dass aufgrund von datenschutzrechtlich gebotenen Maßnahmen im Kontext unzulässiger Verarbeitungen insgesamt ca. $100\text{t}$ CO2-Äquivalent erspart wurden.

## 5.3 Physischer Schutz

Die Maßnahmen 2.C., 3.A.C. und 8.I. sehen (u. a.) einen physischen Schutz des Rechenzentrums vor, d. h. dass unbefugte Personen keinen Zugang zu dem Rechenzentrum erhalten dürfen. Dies kann z. B. durch Mauern, eine Sicherheitsschleuse oder Wachleute umgesetzt werden. Jedoch sind die (physischen Schutz-)Maßnahmen nicht dem Datenschutz zuzurechnen. Es liegt offensichtlich im Interesse der Verantwortlichen, dass Unbefugte keinen physischen Zugang zu den Servern erhalten. Selbst von Verantwortlichen, die bei der

[20] Ada Health GmbH, der Anbieter der bereits zuvor erwähnten Anwendung, die ähnlich zu dem zweiten Teil des Szenarios funktioniert, wollte auf Anfrage keine Informationen zu CO2-Emissionen mitteilen.

Informationssicherheit eher nachlässig sind, dürften es als offensichtlich notwendig erachtet werden, zu verhindern, dass unbefugte Personen Zugang erhalten und damit umfassende und praktisch unbeschränkbare Einwirkungsmöglichkeiten erhalten. Zwar kann die Stringenz der Umsetzung vllt. teilweise dem Datenschutz zugerechnet werden, hierbei handelt es sich jedoch ausschließlich um organisatorische Angelegenheiten, wie die Frage, wer als befugt oder unbefugt gilt oder wie viele Wachleute eingesetzt werden.

Da für die nicht zurechenbaren Teile keine CO2-Emissionen anzurechnen sind und für die organisatorischen Angelegenheiten nach den zuvor benannten Gründen keine CO2-Emissionen zuzuordnen sind, sind für den physischen Schutz vorliegend keine Emissionen zu berechnen.

## 5.4 Abhärtung des Systems

Die Maßnahmen 2.C., 3.A.C. und 8.I. sehen neben den im vorherigen Abschnitt besprochenen physischen Schutzmaßnahmen auch eine Abhärtung des Systems im Sinne von technischen Schutzmaßnahmen vor. Durch technische Maßnahmen wie eine Firewall oder das Monitoring zur Entdeckung potentieller Schadsoftware oder Angriffe ist sicherzustellen, dass nur diejenigen Zugang zu den Serversystemen erhalten bzw. Daten und Prozesse ändern können, die dazu berechtigt sind. Diese Maßnahmen sind dem Datenschutz nur gering zuzurechnen. Abhärtungswerkzeuge sind bei vielen Systemen vorinstalliert (Beispiel: `iptables` auf Linux-Systemen) bzw. leicht aufzusetzen. Es dürfte auch von eher nachlässigen Verantwortlichen noch als notwendig und interessengerecht anerkannt werden, zu verhindern, dass Unbefugte auf technischer Ebene Zugriff erhalten. Anders als der physische Schutz dürfte hier jedoch bei Verantwortlichen mehr Nachlässigkeit zu erwarten sein, wie stringent die Maßnahmen umgesetzt werden. Bereitgestellte Firewall-Tools müssen auch richtig konfiguriert werden, Alerts, die von dem Monitoring-Tool produziert werden, müssen auch organisatorisch angemessen überprüft werden. Insofern verlangt der Datenschutz eine überdurchschnittliche Konsequenz, die ihm – gering – anzurechnen ist. Hierbei handelt es sich auch nicht um ausschließlich organisatorische Maßnahmen, da strengere Regeln eine erhöhte Bearbeitungszeit bedingen [25].

Nach Hayajneh u.a. soll der Rechenmehraufwand einer Firewall durchschnittlich zwischen $200\text{ms}$ bis $600\text{ms}$ [14] betragen. Nach Melkov u.a. erhöht sich die Bearbeitungszeit mit einer Firewall für die Anfrage durchschnittlich bis auf das Doppelte [25]. In Übereinstimmung dieser Wertangaben kann für die Zwecke dieser Arbeit angenommen werden, dass eine Firewall pro Anfrage einen Mehraufwand von ca. $200\text{ms}$ erzeugt. Hiervon sind dem Datenschutz ca. $66.67\text{ms}$ oder ca. $2.23\text{mg}$ CO2-Äquivalent pro Anfrage anzurechnen.

Hierzu kommt der Rechenmehraufwand für weitere Abhärtungsmaßnahmen, wie bspw. ein Monitoring. Heward u.a. geben an, dass verschiedene Monitoring-Methoden im Durchschnitt einen zusätzlichen Aufwand von ca. $40\%$ erzeugen [15]. Dies zugrundegelegt können für die weiteren Abhärtungsmaßnahmen *insgesamt* zusätzliche $50\%$ des Grundaufwands, d.h. zusätzliche $100\text{ms}$ pauschal angenommen werden, von denen dem Datenschutz ca. $33.33\text{ms}$ oder ca. $1.12\text{mg}$ CO2-Äquivalent pro Anfrage anzurechnen sind.

Damit ergibt sich für die Abhärtung des Systems für jede Anfrage ein anzurechnender Mehraufwand von $100\text{ms}$ bzw. $3.35\text{mg}$ CO2-Äquivalent; folglich gilt $\Delta E_{\text{HTTP}} \mathrel{+}= 3.35\text{mg}$.

## 5.5 Redundanz und Wiederherstellbarkeit

Die technisch-organisatorischen Maßnahmen 2.F., 4.B., 5.C. und 8.K. sehen die Redundanz, vorrangig von Prozessen, und ihre Wiederherstellbarkeit bei der Terminbuchung, der Terminerinnerung, der Vor-Ort-Identifizierung und der Konsultation des Chatbots vor. Sie sind im Allgemeinen nur gering zuzurechnen (zu $1/3$), da die Maßnahmen ganz überwiegend (auch) im Interesse der Verantwortlichen liegen, da ein Systemausfall ein erhebliches wirtschaftliches Risiko ist. Jedoch ist ein geringes Maß an Zurechnung geboten, da die vom Datenschutzrecht gebotene Konsequenz und Kontrolle über das offensichtlich Notwendige hinausgeht. Zudem ist für die Maßnahme 4.B. eine hohe Zurechnung (zu $2/3$) anzunehmen, da es nicht unwahrscheinlich ist, dass Verantwortliche das Risiko eines Ausfalls einer Terminerinnerung als geringfügig einschätzen und somit die Maßnahme hauptsächlich durch das Datenschutzrecht motiviert ist.

Diese Maßnahmen führen auf technischer Ebene dazu, dass in Bezug auf die Verarbeitungsvorgänge 2, 4, 5 und 8 die Prozesse doppelt ausgeführt und die Daten doppelt gespeichert werden müssen. Dies führt nicht nur zu einer Verdoppelung des Speichers, sondern auch zu einer Verdoppelung des Berechnungsaufwands, da die Kapazität für eine effektive Wiederherstellbarkeit frei gehalten werden muss und somit nicht von anderen Programmen beansprucht werden kann.

Diese Erhöhung um $100\%$ ist jedoch nur zu $1/3$, bzw. bei 4.B. zu $2/3$ zuzurechnen, sodass dem Datenschutz Erhöhungen von $33.33...\%$ bzw. $66.66...\%$ zuzurechnen sind. Insofern ist der Faktor $\Delta N_x \mathrel{*}= 1.333...$ bzw. $\Delta N_x \mathrel{*}= 1.666...$ anzunehmen. Dieser erfolgt bei der Maßnahme 2.F. für alle drei Nutzungsszenarien bei der Terminbuchung (I., II. und III.) bzgl. der Operationen HTTP-Anfragen, DB-Abfragen und Speicher, bei der Maßnahme 4.B. für das Nutzungsszenario V., bei der Maßnahme 5.C. für das Nutzungsszenario VI.; sowie bei der Maßnahme 8.K. für die Nutzungsszenarien mit Chatbot (II. und III.) bzgl. der Chatbot-Abfragen.

## 5.6 Gewährleistung durch Software/Konfiguration

Den Maßnahmen 4.C., 8.F., 8.L. und 8.O. ist gemeinsam, dass sie die Gewährleistung von Datenschutzprinzipien durch Software oder durch Konfiguration vorsehen und dies mit einer organisatorischen Komponente verbunden ist, durch die es den Verantwortlichen verboten wird, von diesem Softwareablauf oder dieser Konfiguration abzuweichen. Bei Maßnahme 4.C. hat die Software durch Festlegung und Kontrolle der Nutzung (z. B. Double-Opt-In-Verfahren) zugelassener Kommunikationskanäle zu gewährleisten, dass Erinnerungen ausschließlich an die richtigen E-Mail-Adressen gesendet werden. Bei 8.F. ist die Software so zu gestalten, dass in den Chatbot eingegebene Gesundheitsinformationen nicht länger als notwendig gespeichert werden. Bei 8.L. ist durch Voreinstellungen und eine entsprechende Auswahl der Trainingsdaten zu gewährleisten, dass der Chatbot nur die unbe-

dingt notwendigen Gesundheitsinformationen abfragt. Bei 8.O. ist zu gewährleisten, dass die Terminsbuchung auch ohne Nutzung des Chatbots möglich ist. Von diesen Maßnahmen ist Maßnahme 4.C. dem Datenschutz hoch zuzurechnen, da es verständige Verantwortliche zwar für sinnvoll halten dürften, dass E-Mails nur an die richtigen Empfängerinnen und Empfänger gehen, jedoch in der Regel das spezifische Risiko von Fehlschüssen vorrangig in der (hier ggf. als vernachlässigbar angesehene) Verfügbarkeit und weniger in der Vertraulichkeit oder Nichtverkettung sehen, sodass diese Maßnahme überwiegend durch den Datenschutz motiviert ist. Die weiteren Maßnahmen sind voll zuzurechnen, da sie in Widerspruch zu dem Interesse der Verantwortlichen stehen, sich eine „Hintertür" ggf. für zukünftige Nutzungszwecke offenzuhalten.

Die Maßnahmen 8.F., 8.L. und 8.O. zeichnen sich dadurch aus, dass die Gewährleistung vorrangig auf organisatorischer Ebene, nämlich durch Unterlassung der Einbindung bestimmter Funktionen geschieht. Bei Maßnahme 8.O. bspw. ist fortsetzend zu gewährleisten, dass die Nutzung der Terminbuchungssoftware auch ohne Inanspruchnahme des Chatbots möglich bleibt. Hierdurch fällt differentiell jedoch kein Mehr- oder Minderaufwand an, da dies in dem Ausgangssystem zurzeit möglich ist. Das gleiche gilt bei Maßnahme 8.F., denn eine (längerfristigere) Speicherung der Eingaben ist zurzeit nicht vorgesehen. Zwar verlangt Maßnahme 8.L. ggf. technische Änderungen, insbesondere die Anpassung der Logik des Chatbots bzw. der so genannten „Prompts", jedoch kann daraus nicht notwendigerweise auf einen Mehr- oder Minderaufwand geschlossen werden, denn der Berechnungsaufwand für Chatbot-Abfragen hängt, wie oben beschrieben, im Wesentlichen von der Ausgabelänge ab; auf diese nehmen die in Maßnahme 8.L. beschriebenen Konfigurationsanpassungen nur sehr begrenzt Einfluss. Mithin kann für diese drei Maßnahmen keine modellierbare Veränderung festgestellt werden.

Die Maßnahme 4.C. verlangt einerseits die Gewährleistung, dass die in dem System eingespeicherte E-Mail-Adresse richtig ausgelesen und verwendet wird, andererseits auch, dass gewährleistet wird, dass die E-Mail-Adresse auch richtig ist. Sofern die Aktualität betroffen ist, ist insoweit der Maßnahme 4.D. und soweit die Überprüfung der richtigen Auslesung und Verwendung durch Software-Tests der Maßnahme 4.A. Zurechnungsvorrang zu geben. Hier verbleibt folglich noch die Untersuchung des Double-Opt-In-Verfahrens. Dieses ist als neues Nutzungsszenario zu konstruieren:

> **VII. Double Opt-In-Verfahren:** Dieser Use Case umfasst die Verifikation und Bestätigung der Nutzung einer E-Mail-Adresse. Er läuft wie folgt ab:
>
> 1. Anzeige einer Maske für die Eingabe einer E-Mail-Adresse (1 HTTP-Anfrage)
> 2. Eintragung der E-Mail-Adresse als „vorläufig" und Versenden einer E-Mail (1 HTTP-Anfrage, 1 Datenbank-Abfrage, 1 E-Mail)
> 3. Klick auf den Verifizierungslink in der E-Mail und Anzeige einer Bestätigungsseite (1 HTTP-Anfrage, 1 Datenbank-Abfrage)
>
> Damit umfasst der Use Case 3 HTTP-Anfragen, 2 Datenbank-Abfragen und die Versendung einer E-Mail.

Dieses Nutzungsszenario wird wenigstens immer dann ausgeführt, wenn ein neues Benutzerkonto angelegt wird. Nach den oben genannten Annahmen werden jährlich 4382920

Benutzerkonten verwendet. Hiervon sind aber nur diejenigen Benutzerkonten zu berücksichtigen, die neu angelegt werden. Deutlich überschätzend könnte man unterstellen, dass die Zahl der Personen, die in einem Jahr die Plattform nutzen, unabhängig von dem bisherigen Bestand eines Benutzerkontos ist. Daraus ergäbe sich, dass nur $18.5\%$ der im Vorjahr bestehenden Benutzerkonten weiter verwendet würden und somit $81.5\%$ dieser Benutzerkonten oder ca. $3572079$ Benutzerkonten jährlich neu angelegt werden. Dies entspricht aber nicht dem zu erwarteten Verhalten, denn Personen, die bereits ein Benutzerkonto haben, dürften eine deutlich höhere Wahrscheinlichkeit haben, dieses Benutzerkonto auch in Zukunft zu verwenden. Hierfür spricht z. B., dass sie Patient*innen bei Arztpraxen sind, die die Plattform nutzen sowie ihre offenbare Bereitschaft, die Plattform zu verwenden. Realistischer erscheint daher eine Schätzung, dass nur vllt. $20\%$ der Benutzerkonten jährlich neu angelegt würden. Dies wären dann $876584$ Benutzerkonten.

Somit ergeben sich für diesen Use Case pro Jahr $2629752$ HTTP-Anfragen, $1753168$ Datenbank-Abfragen sowie die Versendung von $876584$ E-Mails.

## 5.7 Änderung der Benutzeroberfläche

Die Maßnahmen 1.A., 1.D., 2.G., 2.J., 4.E., 4.F., 5.B., 5.J., 8.A. und 8.N. sehen Änderungen an der Benutzeroberfläche bzw. an den Prozessabläufen bei der Verwendung der Plattform vor. Von ihnen sind die Maßnahmen 1.A. (Möglichkeit zum Ändern oder Löschen von eingetragenen Daten) und 2.J. (Möglichkeit der Terminänderung) dem Datenschutz nicht zuzurechnen. Die Möglichkeit der Aktualisierung der Arztprofile dürfte von den Verantwortlichen offensichtlich für notwendig gehalten werden. Die Möglichkeit der Terminsänderung wird durch Verarbeitungsvorgang 3 bereits nach dem Konzept der Verantwortlichen umgesetzt, sodass auch diese dem Datenschutz nicht zuzurechnen ist. Maßnahme 5.B. (Gewährleistung alternativer Möglichkeiten zur Authentifizierung vor Ort und zur Wiederherstellung des Buchungscodes) ist dem Datenschutz hoch, d.h. zu $2/3$ zuzurechnen, da zwar verständige Verantwortliche eine solche Maßnahme in der Regel ohne datenschutzrechtliche Pflicht für vorteilhaft halten würden, jedoch das Ausmaß der Maßnahme stark datenschutzrechtlich motiviert ist. Die übrigen Maßnahmen sind alle dem Datenschutz voll zuzurechnen, da sie interessenswidrig Beschränkungen oder die Gewährleistung von Freiwilligkeit vorsehen.

Bei den (zumindest teilweise) zuzurechnenden Maßnahmen ergeben sich die folgenden emissionsrelevanten Auswirkungen:

Maßnahme 5.B. besteht aus zwei Bestandteilen, einerseits einer Erweiterung der Authentifizierungsabfrage, um die Suche nach anderen Kriterien als dem Buchungscode sowie die Ergänzung eines Verfahrens zum Zurücksetzen bzw. Wiederherstellen des Buchungscodes. Die Erweiterung der Abfrage dürfte keine wesentlichen Auswirkungen haben, da nur das Feld geändert wird, das durchsucht wird. Für das zusätzliche Verfahren ist ein neues Nutzungsszenario zu konstruieren:

> **VIII. Wiederherstellung eines Buchungscodes:** Dieser Use Case umfasst die Wiederherstellung des Buchungscodes für einen Termin. Da der Buchungscode auf der Terminseite angezeigt wird, entspricht der Use Case den Schritten

1. bis 3. des Use Cases IV, was insgesamt 5 HTTP-Anfragen und 3 Datenbank-Abfragen umfasst.

Nimmt man an, dass dieser Use Case in nur $5\%$ aller Arzttermine verwendet wird, bedeutet die Maßnahme also weitere $7505$ HTTP-Anfragen und weitere $4503$ Datenbank-Abfragen pro Tag, unter Berücksichtigung der Anrechnung von nur $2/3$, ergibt dies einen anrechenbaren Zusatzaufwand von ca. $5003$ HTTP-Anfragen und ca. $3002$ Datenbank-Abfragen.

Die Maßnahmen 5.J. und 8.N. sind unter den Maßnahmen zur Änderung der Benutzeroberfläche solche, bei denen die Änderung im Wesentlichen nur im Austausch einzelner Daten oder Datenfelder ohne messbare Änderung bei den HTTP-Anfragen, den Datenbank-Abfragen, den Chatbot-Abfragen oder im Speicher besteht. Bei Maßnahme 5.J. ist dies offensichtlich, da nur einzelne Felder durch Pseudonyme ersetzt werden sollen und deren Berechnung keinen großen Aufwand erfordern dürfte. Bei Maßnahme 8.N. kann zwar der Eindruck entstehen, dass ergänzende Datenfelder vorzusehen sind, was einen Speicheraufwand bedeutet, jedoch sieht die Maßnahme dies nicht ausdrücklich vor. Ein solches Speicherfeld kann nämlich auch z. B. als URL-Parameter implementiert werden, sodass ein zusätzlicher Speicheraufwand nicht zwingend ist. Soweit zusätzlicher Rechenaufwand durch z. B. eine zwischengeschaltete Seite entstehen könnte, dürfte dieser eher der Maßnahme 8.A. zugerechnet werden und wird daher mit dieser behandelt, um eine doppelte Bilanzierung zu vermeiden.

Die Maßnahme 8.A. verlangt, dass eine Warnung vor der Verwendung des Chatbots eingeblendet wird. Dies bedeutet, dass bei den Use Cases II und III jeweils zwischen dem 2. und 3. Schritt eine zusätzliche Informationsseite eingeblendet wird, was mit einer zusätzlichen HTTP-Anfrage verbunden ist, d.h. $\Delta s_{\mathrm{II,HTTP}} \mathrel{+}= 1$ und $\Delta s_{\mathrm{III,HTTP}} \mathrel{+}= 1$. Auf dieser Informationsseite könnten dann auch die Maßnahmen 8.N. und 8.O. umgesetzt werden.

Die Maßnahmen 1.D. und 2.G. verlangen, dass die (standardmäßig) erhobenen Daten bei den Arztprofilen bzw. bei Terminbuchungen auf das unbedingt notwendige beschränkt werden. Sie stehen damit jedoch einer freiwilligen Erhebung nicht entgegen, solange die Freiwilligkeit eindeutig dargestellt ist, tatsächlich gegeben ist und keine Nachteile aus der Nichterhebung oder der Löschung der freiwillig erhobenen Daten entstehen. Mit diesen Auflagen können die Daten auch gespeichert werden, sodass Datenfelder bereitgehalten werden können. Da in festen Datenstrukturen auch leere Datenfelder Speicherplatz belegen, kann keine Reduktion der gespeicherten Daten ermittelt werden. Die Auflagen bestehen nur in der Einfügung von Informationstexten neben den Feldern oder in organisatorischen Weisungen, denen ebenfalls keine Klimafolgen zugewiesen werden.

Nach Maßnahme 4.E. sollen die in der Erinnerungs-E-Mail enthaltenen Informationen zu dem Termin und der betroffenen Person auf ein notwendiges Minimum reduziert werden. Hieraus folgt jedoch nicht notwendigerweise eine Veränderung bei den Datenbankabfragen, denn es dürfte weiterhin erforderlich sein – in einer oder mehreren Abfragen – die Person (inbs. E-Mail-Adresse), den Termin (insb. Datum) und die Arztpraxis (insb. Bezeichnung, ggf. Adresse) abzufragen. Eine Reduktion bei diesen vier Datenpunkten würde das notwendige Minimum unterschreiten, da entweder die E-Mail mangels Zieladresse nicht versendet, oder eine E-Mail, die den Zeitpunkt oder Ort des Termins nicht angibt,

nicht mehr als noch zweckmäßig angesehen werden könnte. Die in Nutzungsszenario V bereits vorgesehenen Abfragen entsprechen jedoch diesem Minimum, sodass hier nichts abzuschlagen ist. Zu beachten ist auch, dass eine Abfrage potentiell mehr Daten zurückgeben kann, als am Ende in der E-Mail dargestellt oder verwendet werden. Der Berechnungsaufwand für den Zusammenbau von E-Mails aus einer Vorlage dürfte sich nur geringfügigst in einem nicht messbaren Ausmaß ändern.

Maßnahme 4.F. sieht ein Speicherfeld vor, in dem für jeden gebuchten Termin zusätzlich ein Wahrheitswert gespeichert wird, nämlich ob eine Terminerinnerung gewünscht ist, sowie einen Use Case zur Bearbeitung dieses Speicherfelds. Wahrheitswerte werden, auch wenn formal ein Bit genügt, üblicherweise aus Gründen der Speichervereinheitlichung in wenigstens einem Byte gespeichert. [21] Dieses Byte ist den Workflows I und II hinzuzufügen, d.h. $\Delta s_{\mathrm{I,Speicher}} \mathrel{+}= 1\mathrm{B}$ und $\Delta s_{\mathrm{II,Speicher}} \mathrel{+}= 1\mathrm{B}$. Da diese für $80\%$ aller Buchungsfälle angewendet werden, entspricht diese Maßnahme für sich gestellt einem zusätzlichen täglichen Speicheraufwand von ca. $24\mathrm{KB}$.

Zudem ist ein weiterer Use Case für die Bearbeitung des Speicherfelds erforderlich:

> **IX. Änderung der Auswahl einer Terminerinnerung:** Dieser Use Case umfasst die Änderung der Auswahl, ob eine Terminerinnerung gewünscht wird. Da die Änderung auf der Terminseite erfolgen kann und hierfür keine zusätzlichen Informationen erforderlich sind, bedarf es keiner Zwischenseite, somit entspricht der Use Case den Schritten 1. bis 3. und 5. des Use Cases IV, was insgesamt 6 HTTP-Anfragen und 4 Datenbank-Abfragen umfasst.

Es ist davon auszugehen, dass dieser Use Case nur in einer verhältnismäßig geringen Anzahl von Fällen ausgeführt wird, da bereits mit der Terminbuchung angegeben werden kann, ob eine Terminerinnerung gewünscht wird. Nimmt man darauf beruhend an, dass dieser Use Case in $0.5\%$ aller Arzttermine verwendet wird, bedeutet die Maßnahme also weitere $900$ HTTP-Anfragen und weitere $600$ Datenbank-Abfragen pro Tag.

## 5.8 Regelmäßige Abfragen

Den Maßnahmen 1.B., 4.D., 8.B. und 8.H. ist gemein, dass sie die Festlegung und Umsetzung eines Konzepts Prozessen zur Löschung, Berichtigung oder Aktualisierung von personenbezogenen Daten vorsehen. Diese Prozesse sind durch (z. B. mittels `cronjob`) regelmäßig ausgeführte Abfragen umzusetzen, die unter allen Daten anhand bestimmter Kriterien diejenigen herausfiltern, die falsch, veraltet oder löschreif sind und anschließend

[21] Dies kann man z.B. für die Programmiersprache C leicht selber überprüfen mit diesem Programm:

```c
#include <stdio.h>
#include <stdbool.h>

int main() {
    bool send_email = true
    printf("%u", sizeof(send_email) * 8);
}
```

Bei Ausführung gibt es eine `1` aus und verdeutlicht damit, dass das der Datentyp eine Größe von einem Byte hat, da `sizeof` nach § 6.5.3.4. der C-Spezifikation die Größe in Bytes angibt [17].

die jeweils notwendige Handlung vornehmen. Von diesen Maßnahmen sind die Maßnahmen 1.B. und 8.B. dem Datenschutz gering (zu $1/3$) zuzurechnen, da es ganz überwiegend auch im Interesse der Verantwortlichen liegt, die Angaben zu den Arztpraxen bzw. die Trainingsdaten für den Chatbot aktuell und richtig zu halten, da andernfalls die Plattform nicht so effektiv, wie beabsichtigt, angeboten würde und mit einem Abgang von Nutzenden zu rechnen wäre; unbeschadet dessen ist die konkrete Zielrichtung und die Konsequenz der Maßnahmen weiterhin zumindest teilweise durch den Datenschutz motiviert, sodass die Zurechnung zumindest teilweise verbleibt. Die Maßnahmen 4.D. und 8.H. sind dem Datenschutz voll zuzurechnen, da sie sich überwiegend in Konflikt zu den Interessen der Verantwortlichen begeben, indem die Löschung von Daten vorgeschrieben wird, an deren weiterer Speicherung die Verantwortlichen möglicherweise noch ein Interesse hätte. Hieran kann auch bei Maßnahme 4.D. die Verpflichtung, die E-Mail-Adressen aktuell zu halten, die für sich alleine wegen überwiegender Motivation durch den Datenschutz jedenfalls hoch zuzurechnen wäre, nichts ändern.

Die Maßnahme 8.B. ist vorrangig organisatorisch, da die Löschung und Berichtigung falscher Trainingsdaten am Ende nur manuell vorgenommen werden kann und insoweit die Maßnahme durch organisatorische Prozesse, die eine regelmäßige manuelle Überprüfung von Trainingsdaten und ggf. eine Bestätigung oder ein Peer Review vorsehen. Insoweit können der Maßnahme keine modellierbaren Emissionen zugeordnet werden.

Die Maßnahme 1.B. kann umgesetzt werden, indem regelmäßig (z. B. einmal pro Jahr) per E-Mail daran erinnert wird, die Informationen zur Arztpraxis aktuell zu halten. Eine solche „Selbstkontrolle“ dürfte die geeignetste Umsetzungsform sein, da von Seiten der Plattform die Richtigkeit und Aktualität nur schwer bzw. mit deutlich eingriffsintensiveren Mitteln (bspw. umfassendes Web-Scraping) zu erreichen versucht werden kann. Somit sind nach den oben beschriebenen Annahmen zur Verbreitung jährlich $8190$ E-Mails zu senden. Diese sind jedoch nur zu $1/3$ zuzurechnen, sodass dem Datenschutz $2730$ zusätzliche E-Mails.

Die Maßnahme 4.D. kann durch regelmäßige Popups bzw. Alerts umgesetzt werden. Eine Aufforderung zur Aktualisierung per E-Mail dürfte nicht erforderlich sein, da E-Mail-Erinnerungen nur aufgrund einer expliziten Terminbuchung entstehen können, sodass alle erforderlichen Informationen bei dieser abgefragt werden können. Nimmt man, wie bei 1.B. an, dass eine Abfrage pro Jahr genügt, wird bei $4382920$ Personen eine zusätzliche Zwischenseite nach der Anmeldung erforderlich; diese ist jeweils mit einer zusätzlichen HTTP-Anfrage verbunden, die voll anzurechnen ist.

Die Maßnahme 8.H. kann durch eine regelmäßige Löschroutine mit relativer Festlegung des Abschnittzeitpunkts, ähnlich dem Prinzip eines *rolling windows*, erfolgen. Dies führt dazu, dass der Umfang der zu löschenden Daten im Wesentlichen stets gleich bleibt und pauschalisiert dem Anfall neuer Daten in dem Zeitraum zwischen zwei Ausführungen des Löschprozesses entspricht. Bei täglicher Ausführung sind damit Daten in dem Umfang zu löschen, wie sie täglich aufgrund der Nutzung des Chatbots anfallen. Nach den oben beschriebenen Annahmen sind dies $90\%$ der täglichen $30020$ Terminbuchungen bzw. $27018$ Chatbot-Nutzungen. Theoretisch kann die Löschung dieser Eintragungen in einer einzi-

gen Datenbank-Abfrage[22] erfolgen. Dies würde jedoch verzerren, dass eine solche Abfrage deutlich aufwändiger ist als die Löschung bspw. nur eines einzigen Eintrags. Daher ist es sinnvoller, den Vorgang wie bei der Beschreibung von Use Case V. durch eine allgemeine Abfrage zur Bestimmung der betroffenen Einträge und je Eintrag eine weitere Abfrage zur Löschung dessen zu modellieren. Dies ergibt $27019$ zusätzliche tägliche Datenbank-Abfragen.

## 5.9 Authentifizierung und Autorisierung

Die Maßnahmen 2.A. und 3.A.A. sehen die Authentifizierung und Autorisierung als Voraussetzung für den Zugriff auf Termindaten vor. Diese Maßnahmen sind dem Datenschutz höchstens gering zuzurechnen, da durchschnittliche, verständige Verantwortliche diese Maßnahme in der Regel auch ohne gesetzliche Pflicht für sehr vorteilhaft, wenn nicht gar notwendig halten dürften. Die Beschränkung der Zugriffsmöglichkeiten auf die Termine und Terminsänderung auf solche Personen, die sich identifiziert haben und nachweislich berechtigt sind, dürfte essentielle Grundlage für eine vertrauensvolle Arzt-Patienten-Beziehung und fundamentaler Gegenstand der ärztlichen Schweigepflicht sein. Zudem ist es offensichtlich auch im Interesse der Verantwortlichen, dass Termine nur von berechtigten Personen geändert werden.

Vor dem Hintergrund dieser Bedeutung und, dass das Verlangen einer Anmeldung nach dem Stand der Technik bei Webplattformen allgemein verbreitet ist, stellt sich die Frage, ob die Maßnahmen nicht sogar ohne Zurechnung bleiben. Diese Frage kann jedoch aus den folgenden Gründen offen gelassen werden:

Die Maßnahme 2.A., soweit sie sich auf den Zugang der Patient*innen bezieht und die Maßnahme 3.A.A. sind bereits in dem System umgesetzt und zwar insbesondere in dem Schritt 1 des Use Cases IV. Insoweit bringt das Datenschutzrecht keine zusätzliche Differenz ein. Auch die Identifizierung des ärztlichen Personals ist bei der Vor-Ort-Identifikation unterstellt. Daher dürfte eine weitere Differenz auch insoweit nicht verbleiben.

## 5.10 Verschlüsselung

Die Maßnahmen 2.D., 3.A.D., 5.H. und 8.G. sehen den Einsatz von Verschlüsselung vor und zwar insbesondere den Einsatz von Transportverschlüsselung. Die Maßnahmen 2.D., 3.A.D. und 5.H. sind dem Datenschutz jeweils höchstens gering zuzurechnen, da Transportverschlüsselung im Internet inzwischen weit verbreitet ist und nach dem Stand der Technik mit minimalem Aufwand umsetzbar ist.

Verschlüsselung könnte sowohl Speicher- als auch Rechenaufwand begründen. Die Ver- und Entschlüsselung kostet Rechenzeit. Zudem könnte Verschlüsselung dazu führen, dass die gespeicherten Daten mehr Platz einnehmen.

[22] Beispielsweise mit der folgenden SQL-Abfrage (MySQL-Dialekt):

```
DELETE FROM chatbot_requests WHERE created_at <= DATE_SUB(NOW(), INTERVAL -1 YEAR)
```

Der Rechenaufwand usw. hängt maßgeblich von der Menge der zu verschlüsselnden Daten ab. Eine gute Annahme der Größe durchschnittlicher Webseiten liegt zwischen $25\mathrm{KB}$ und $200\mathrm{KB}$. In diesem Bereich dürfte ein Rechenaufwand von ca. $100\mathrm{ms}$ bis $150\mathrm{ms}$ anzunehmen sein [13, 2]. Von diesen $150\mathrm{ms}$ sind dem Datenschutz nur $1/3$ zuzurechnen, d.h. ein Rechenaufwand von $50\mathrm{ms}$. Insofern gibt es daher bei HTTP-Anfragen eine zusätzliche Emission von $\Delta E_{\mathrm{HTTP}} \mathrel{+}= 1.675\mathrm{mg}$ zurechenbarem CO2-Äquivalent.

Transportverschlüsselung verwendet in der Regel hybride Verschlüsselung. Insofern ergeben sich bei dem Speicheraufwand die folgenden Beobachtungen: Moderne symmetrische Verschlüsselungsverfahren erhöhen den Speicheraufwand in der Regel nicht unmittelbar, sondern nur durch zwei Modifikationen: einerseits fügen Verschlüsselungssysteme häufig zusätzliche Informationen an, wie bspw. eine PGP-Armor oder durch Angabe einer HMAC. Andererseits sind moderne symmetrische Verschlüsselungsverfahren wie AES in der Regel so genannte Block-Chiffren; lässt sich der Klartext nun nicht unmittelbar in eine ganze Anzahl von Blöcken zerteilen, wird ein Padding-Verfahren verwendet, um alle Blöcke aufzufüllen. Die erste Modifikation hat eine im Verhältnis zur Datenmenge feste Länge und die zweite Modifikation ist durch die Blockgröße nach oben beschränkt, sodass beide Modifikationen asymptotisch vernachlässigbar sind. Das gleiche gilt für den Teil der symmetrischen Verschlüsselung, der im Wesentlichen der Übermittlung des Schlüssels für den symmatrischen Teil dient.

## 5.11 Trennung

Die Maßnahmen 2.E., 3.A.E. und 5.I. sehen die logische und ggf. physische Trennung von Daten vor. Trennung verhindert oder hemmt, dass Daten versehentlich oder unberechtigt verkettet werden können, und unterstützt Zugriffsberechtigungssysteme. Die Maßnahmen sind dem Datenschutz voll zuzurechnen, da sie ein hohes Maß an technischer und organisatorischer Komplexität begründen, die von verständigen Verantwortlichen in der Regel nicht für unbedingt erforderlich gehalten werden dürfte; zudem verhindert die Trennung Verkettungs- und Zugangsmöglichkeiten und steht damit zumindest teilweise im Widerspruch zum Interesse der Verantwortlichen.

Jedoch bringt die Trennung keine modellierbaren Klimafolgen mit sich. Zwar führt eine Trennung – wie Redundanz – auch dazu, dass Prozesse mehrfach ausgeführt oder Systeme mehrfach bereitgehalten werden müssen. Anders als die Redundanz verändert sich jedoch der Gesamtaufwand nicht, denn es muss gerade kein Berechnungsaufwand in entsprechender Höhe „freigehalten“ werden, sondern der Aufwand verteilt sich nur auf mehrere Einzelsysteme.

Bei 5.I. liegt zudem vorrangig eine organisatorische Maßnahme vor, denn die Trennung soll hier v. a. an dem Zugangscomputer in der Praxis erfolgen und kann so durch mehrere Benutzerkonten und ein Berechtigungskonzept organisatorisch ohne technische Änderungen umgesetzt werden.

## 5.12 Protokollierung

Die Maßnahmen 2.H., 3.C. und 5.E. sehen die Protokollierung von Zugriffen und Änderungen sowie deren Überwachung in Bezug auf Termine und Terminsänderungen bzw. im Rahmen der Vor-Ort-Identifikation vor. Sie sind dem Datenschutz hoch (zu $2/3$) zuzurechnen, da zwar Verantwortliche in der Regel eine gewisse Protokollierung vornehmen werden (z. B. von HTTP-Anfragen oder von Datenbank-Abfragen), jedoch die Maßnahmen einen etwas anderen Fokus legen, der vorrangig durch den Datenschutz motiviert ist.

Alle drei Maßnahmen betreffen die selbe Tätigkeit, nämlich den Zugriff auf Terminsdaten sowie die damit verbundenen Daten über die Patient*innen. Daher liegt es nahe, ein gemeinsames Protokoll über die Tätigkeit bei allen drei Verarbeitungsvorgängen zu einzusetzen. Dieses müsste speichern, wer auf welchen Termin wann zugreift und, wenn ein Grund eingegeben werden muss oder durch das System erkannt werden kann, warum der Zugriff erfolgt. Zudem sollte protokolliert werden, welche Daten konkret angefragt werden, also insbesondere, ob auch die bei der Buchung eingegebenen Gesundheitsdaten wie bspw. das jeweilige Leiden abgefragt werden. Bei der Änderung eines Termins ist zudem zu protokollieren, welche Änderung vorgenommen wurde.

Der vorrangige Faktor bei der Protokollierung ist der Speicheraufwand. Nimmt man an, dass die verantwortende Person und der Termin jeweils durch eine ID-Nummer üblicherweise im Umfang eines 64-Bit-Integers (8 Byte) repräsentiert werden und für das Datum das ISO 8601-Format verwendet wird, welches für eine Datums- und Zeitangabe zur UTC-Zeit mindestens 17 Zeichen[23] verlangt, die in UTF8-Codierung mit je einem Byte dargestellt werden können. Die Angabe eines Grundes kann sehr viel oder sehr wenig Speicherplatz in Anspruch nehmen. Es sei daher unterstellt, dass nur eine knappe Angabe erforderlich gemacht wird, die wegen Auswahlfeldern in der Praxis selten manuell befüllt wird. In diesem Fall kann man annehmen, dass 128 Byte (weniger als 128 Zeichen) ausreichen dürften. Die Protokollierung der Abfrage auch der eingegebenen Gesundheitsdaten kann mit einem einzigen Byte erfolgen. Für die Beschreibung der Änderung können – durchschnittlich – weitere 128 Byte veranschlagt werden, wenn man annimmt, dass die Änderung mehrerer Felder durch mehrere Protokolleinträge dargestellt wird. Damit ergeben sich 290 Byte für einen Protokolleintrag.

Protokolleinträge werden angelegt, wenn der Use Case IV oder VI ausgeführt wird, dies ist täglich in $10\%$ der Termine bzw. bei jedem Termin der Fall, insgesamt werden also täglich ca. $33022$ Protokolleinträge mit insgesamt $9.58\mathrm{MB}$ angelegt. Dies entspricht einer täglichen Emission von $0.535\mathrm{mg}$ CO2, die durch die Protokollierung entsteht. Von dieser sind dem Datenschutz nach vorstehenden Erwägungen $0.356\mathrm{mg}$ anzurechnen.

## 5.13 Software-Tests

Die Maßnahmen 4.A., 5.D. und 8.C. sehen vor, dass das Sollverhalten von Prozessen festgelegt und dokumentiert, d.h. spezifiziert, sowie regelmäßig getestet wird, sodass sicher-

[23] Beispiel: `2023-10-16T14:50:00Z`

gestellt werden kann, dass die Erinnerung an die Termine (4.A.) bzw. die Vor-Ort-Identifikation (5.C) richtig funktioniert und der Chatbot korrekte Antworten liefert (8.C.).

Diese Maßnahmen sind dem Datenschutz nicht zuzurechnen, da Software-Tests – auch in hoher Testdichte – allgemein verbreitet sind und eine hohe Testdichte bei hinreichend professionalisierten Softwareentwicklungsprojekten als *lege artis* angesehen werden kann. Daher dürfte es für verständige Verantwortliche offensichtlich notwendig sein, die datenschutzrechtlich geforderten Tests vorzunehmen. Die Maßnahmen 5.C. und 8.C. liegen zudem vollständig im Interesse der Verantwortlichen, da sie für die Sicherstellung der Funktionalität der Plattform essentiell sind und ein Funktionsversagen bei Vernachlässigung der Tests erhebliche negative wirtschaftliche Auswirkungen haben dürfte. Zuvor wurde zwar gesagt, dass die Bedeutung der Funktionalität der Erinnerung vorrangig datenschutzrechtlich motiviert ist, sodass aus diesem Grund 4.A. als wenigstens gering zuzurechnen sein könnte. Dies ist jedoch hierauf nicht übertragbar, da Software-Tests von den Verantwortlichen schon im Eigeninteresse als notwendig erachtet werden dürften, während bei Maßnahmen wie 4.B. oder 4.C. (Redundanz und Wiederherstellbarkeit bzw. Gewährleistung durch Software/Konfiguration) Verantwortliche geneigt sein könnten, das Risiko eines Ausfalls oder Fehlschlags als akzeptabel hinzunehmen.

# Kapitel 6

# Auswertung und Fazit

Nachdem im vorherigen Kapitel für die einzelnen Datenschutzmaßnahmen im Detail bestimmt wurde, welche Klimaauswirkungen sie mit sich bringen, soll nun abschließend eine Gesamtbetrachtung erfolgen. Dazu werden zunächst die einzelnen Faktoren, für die zuvor wegen der gegenseitigen Abhängigkeiten teilweise noch nicht eine unmittelbare Klimafolge berechnet werden konnte, zusammengerechnet. Dies ergibt dann eine Gesamtemissionsdifferenz, welche anschließend diskutiert und bewertet wird.

## 6.1 Gesamtemissionen

Aufgrund der Analyse der einzelnen Datenschutzmaßnahmen haben sich fünf Möglichkeiten ergeben, wie eine Maßnahme auf die Gesamtemissionen Einfluss haben kann: es können die Kosten je Operation verändert, neue Use Cases ergänzt, die Schritte in einem Use Case angepasst, die Anzahl der Ausführungen eines Use Cases geändert oder zusätzliche Emissionsänderungen angerechnet werden. Diese Einflussmöglichkeiten sollen nun in der genannten Reihenfolge zusammengerechnet werden.

**Kosten je Operation** Durch einzelne Maßnahmenbereiche (z. B. die Abhärtung des Systems oder die Verschlüsselung) wurden die Kosten je Operation verändert. Betroffen war von dieser Veränderung ausschließlich die Operation **HTTP-Anfragen**. Hierbei ergibt sich die folgende Berechnung:

| Bereich | Modifikation | Maßnahmen |
|---|---|---|
| Initiale Annahme | $E_{\mathrm{HTTP}} = 6.7\mathrm{mg}$ | |
| 5.4. Abhärtung des Systems | $\Delta E_{\mathrm{HTTP}} \mathrel{+}= 3.35\mathrm{mg}$ | 2.C., 3.A.C., 8.I. |
| 5.10. Verschlüsselung | $\Delta E_{\mathrm{HTTP}} \mathrel{+}= 1.675\mathrm{mg}$ | 2.C., 3.A.C., 8.I. |

Hiermit ergibt sich eine Gesamtemission je HTTP-Anfrage im Umfang von $11.725\mathrm{mg}$ je HTTP-Anfrage.

**Neue Use Cases** Durch Maßnahmen aus den Bereichen der Gewährleistung durch Software/Konfiguration und der Änderung der Benutzeroberfläche werden drei neue Use Ca-

ses eingeführt, die für jede ihrer Ausführungen zusätzliche HTTP-Anfragen, Datenbanken oder versendete E-Mails verantworten:

| | Bezeichnung | Ausführungen | HTTP | E-Mails | Datenbank |
|---|---|---|---|---|---|
| VII | Double Opt-In-Verfahren | $876584\times$ p. a. | $3\times$ | $1\times$ | $2\times$ |
| VIII | Wiederherstellung eines Buchungscodes | $365494\times$ p. a. | $3\times$ | $1\times$ | $2\times$ |
| IX | Änderung der Auswahl einer Terminserinnerung | $54788\times$ p. a. | $3\times$ | $1\times$ | $2\times$ |

**Operationen je Use Case** Durch den Maßnahmenbereich der Änderung der Benutzeroberfläche wurden einzelnen Use Cases weitere Operationen zugeschlagen, und zwar den Use Cases II und III je eine zusätzliche HTTP-Anfrage und den Use Cases I und II je ein zusätzliches Byte Speicher. Damit ergeben sich folgende Schrittzahlen:

| | HTTP (Diff.) | HTTP | Speicher (Diff.) | Speicher |
|---|---|---|---|---|
| I | n. a. | n. a. | $+1$B | 307B |
| II | $+1$ | 38 | $+1$B | 307B |
| III | $+1$ | 20 | n. a. | n. a. |

**Ausführungen je Use Case** Durch den Maßnahmenbereich der Redundanz und Wiederherstellbarkeit wurde für einzelne Use Cases geändert, wie oft sie insgesamt ausgeführt werden. Den Use Cases I, II, III und VI wurden $33,3\%$ zugeschlagen; dem Use Case V $66,7\%$. Damit ergeben sich die folgenden (insgesamten) jährlichen Ausführungszahlen:

| | Zuvor | Änderung | Ergebnis |
|---|---|---|---|
| I | $1096480\times$ | $+33.3\%$ | $1461973\times$ |
| II | $7675363\times$ | $+33.3\%$ | $10233818\times$ |
| III | $2192961\times$ | $+33.3\%$ | $2923948\times$ |
| IV | $1096481\times$ | $\pm0$ | 1096480 |
| V | 365 | $+66.7\%$ | $609\times$ |
| VI | $10964805\times$ | $+33.3\%$ | $14619740\times$ |
| VII | $876584\times$ | $\pm0$ | $876584\times$ |
| VIII | $365493\times$ | $\pm0$ | $365493\times$ |
| IX | $54787\times$ | $\pm0$ | $54787\times$ |

Mit diesen Ausführungszahlen lässt sich berechnen, wie oft jährlich HTTP-Anfragen anfallen, E-Mails versendet, Abfragen an eine Datenbank oder an einen Chatbot gestellt werden oder Daten in dem angegebenen Umfang gespeichert werden:

| | HTTP | E-Mail | Datenbank | Chatbot | Speicher |
|---|---|---|---|---|---|
| I | 30701433× | 0× | 24853541× | 0× | 335.98GB |
| II | 388885084× | 0× | 173974906× | 133039634× | 2358.89GB |
| III | 58478960× | 0× | 2923948× | 38011324× | 0B |
| IV | 7675360× | 0× | 4385920× | 0× | 0GB |
| V | 0× | 18282180× | 36564969× | 0× | 0B |
| VI | 43859220× | 0× | 29239480× | 0× | 0B |
| VII | 2629752× | 0× | 1753168× | 0× | 0B |
| VIII | 1827465× | 0× | 1096479× | 0× | 0B |
| IX | 328722× | 0× | 219148× | 0× | 0B |

**Emissionen der Use Cases** Mithilfe der vorstehenden Veränderungen lässt sich berechnen, welche Emissionen durch die Anpassung und Ergänzung der Use Cases sowie durch die Veränderung des Rechenaufwands bei HTTP-Anfragen entsteht. Hierzu sind einerseits die vorstehend angegebenen Werte zusammenzurechnen, andererseits die aus den vorstehenden Angaben sowie den Szenariobeschreibungen erkenntliche Ausgangswerte zusammenzurechnen und ihre Differenz zu bilden.

Dadurch erhält man die folgenden jährlichen Gesamtwerte *vor Anwendung der Datenschutzmaßnahmen*:

| | HTTP | E-Mail | Datenbank | Chatbot | Speicher |
|---|---|---|---|---|---|
| Anzahl | 389250552× | 10957300× | 199544791× | 128288212× | 2684.18GB |
| Faktor | 6.7mg | 6.7mg | 25.125mg | 100.5mg | 55.83mg/GB |
| CO2 | 2607.978kg | 73.413kg | 5013.562kg | 12892.965kg | 0.150kg |

Insgesamt entspricht dies einer Gesamtemission der Use Cases von 20.59t CO2-Äquivalent pro Jahr.

*Nach Anwendung der Datenschutzmaßnahmen* erhält man hingegen die folgenden Werte:

| | HTTP | E-Mail | Datenbank | Chatbot | Speicher |
|---|---|---|---|---|---|
| Anzahl | 534386002× | 18282180× | 275011563× | 171050958× | 2695.88GB |
| Faktor | 11.725mg | 6.7mg | 25.125mg | 100.5mg | 55.83mg/GB |
| CO2 | 3657.697kg | 49.076kg | 1896.102kg | 4297.655kg | 0.151kg |

Insgesamt entspricht dies einer Gesamtemission der Use Cases nach Anwendung der Datenschutzmaßnahmen von 30.49t CO2-Äquivalent pro Jahr, einer jährlichen Differenz von +9.9t zu den Emissionen vor Anwendung.

**Zusätzliche Emissionsänderungen** Durch die Maßnahmenbereiche der unzulässigen Verarbeitungen, der regelmäßigen Abfragen und der Protokollierung treten zudem weitere Emissionsänderungen auf, die unmittelbar angerechnet werden müssen. Derartige Änderungen treten mit den Operationen HTTP-Anfragen, E-Mail-Nachrichten, Datenbank-Abfragen, mit Speicheraufwand oder mit sonstigen, unmittelbar bestimmbaren CO2-Emissionen auf.

Für den Maßnahmenbereich der unzulässigen Verarbeitungen ergibt sich:

| Grund | HTTP | E-Mails | Datenbank | Speicher | Sonstige |
|---|---|---|---|---|---|
| Unzulässigkeit VV 6, 7 | | | | | $-100\text{t}$ |

Für den Maßnahmenbereich der regelmäßigen Abfragen ergeben sich:

| Grund | HTTP | E-Mails | Datenbank | Speicher | Sonstige |
|---|---|---|---|---|---|
| TOM 1.B. | | $2730\times$ p. a. | | | |
| TOM 4.D. | $4382920\times$ p. a. | | | | |
| 8.H. | | | $9868690\times$ p. a. | | |

Für den Maßnahmenbereich der Protokollierung ergibt sich:

| Grund | HTTP | E-Mails | Datenbank | Speicher | Sonstige |
|---|---|---|---|---|---|
| TOM 2.H., 3.C., 5.E. | | | | $2332\text{GB}$ p. a. | |

Wendet man hierauf die Emissionsfaktoren für die einzelnen Operationen an, erhält man:

| | HTTP | E-Mails | Datenbank | Speicher | Sonstige |
|---|---|---|---|---|---|
| Summe | $4382920\times$ p. a. | $2730\times$ p. a. | $9868690\times$ p. a. | $2332\text{GB}$ p. a. | $-100\text{t}$ |
| Emissionen | $+51390\text{g}$ p. a. | $+18.29\text{g}$ p. a. | $+247950\text{kg}$ p. a. | $+130.20\text{g}$ p. a. | $-100\text{t}$ |

Es ist daher zu folgern, dass Datenschutzmaßnahmen durch zurechenbare zusätzliche Emissionsänderungen ca. $299.49\text{kg}$ CO2-Äquivalent jährlich zusätzlich verbrauchten und ca. $100000\text{kg}$ CO2-Äquivalent (einmalig) erspart haben.

**Emissionsdifferenz** Unter Verbindung der Emissionsdifferenz bei den Use Cases und der weiteren Emissionsdifferenzen sind dem Datenschutz jährliche Emissionen in Höhe von ca. $10.20\text{t}$ CO2-Äquivalent sowie eine einmalige Einsparung in Höhe von ca. $100\text{t}$ zuzurechnen.

## 6.2 Ergebnisdiskussion

Auf einer rein rechnerischen Ebene ist das Ergebnis dieser Untersuchung klar. Datenschutz spart das zehnfache seiner Mehremissionen ein; die Mehremissionen überholen erst nach zehn Jahren seine Einsparungen. Bedenkt man, dass für Serverhardware durchschnittlich eine Lebensdauer von ca. fünf Jahren eingeplant wird, wonach eine vollständige Erneuerung mit den entsprechenden, nicht unbeachtlichen CO2-Emissionen verbunden ist, fällt Datenschutz somit nur untergeordnet ins Gewicht.

In Bezug auf die Plausibilität dieses Ergebnisses müssen jedoch einige Punkte offenbleiben:

**Unzulässige Verarbeitungen** Der im Rahmen der verwendeten Modellierung und Untersuchung einzige sowie im Ergebnis ausschlaggebende Teil, bei dem Datenschutzmaßnahmen unmittelbar zu einer Verringerung der Emissionen geführt haben, ist die Unzulässigkeit der Verarbeitungsvorgänge 6 und 7 sowie die daraus folgende Unterlassung der Anfertigung des Chatbot-„KI“-Modells. Gegen die Notwendigkeit dieser Datenschutzmaßnah-

me in dem konkreten behandelten Szenario dürfte zwar aus den angegebenen Gründen wenig zu erinnern sein, jedoch handelt es sich bei dem Training eines „KI“-Modells um einen besonders berechnungsintensiven Vorgang. Insoweit stellt sich die Frage, ob und inwieweit dieser Abzug repräsentativ und somit übertragbar ist.

Zwar lässt sich in gewisser Hinsicht eine Korrelation herstellen, wonach berechnungsintensivere Datenverarbeitungsverfahren auch mit erhöhter Wahrscheinlichkeit inhärente Datenschutzprobleme aufweisen (bspw. „KI“, Big Data oder „Blockchain“-Technologien). Dies ist jedoch nur begrenzt, einige spezfische Technologien betreffend, pauschalisierbar.

**Relative Änderung** Verschwiegen werden darf aber auch nicht, dass die Klimabilanz des Datenschutzes aufgrund der Ergebnisse dieser Arbeit zwar grundsätzlich positiv ist, jedoch die zusätzlichen Emissionen (bei Außerachtlassung der unzulässigen Verarbeitungen) relativ gesehen gegenüber den bisherigen Emissionen erheblich ist, nämlich eine Steigerung um gut $50\%$ darstellt (von ca. $21\mathrm{t}$ zu ca. $30\mathrm{t}$).

**Wesentliche Faktoren** Betrachtet man die Auswirkungen der Veränderungen bei den einzelnen Operationen, so ist festzustellen, dass die Steigerung der Chatbot-Anfragen insgesamt gut $43\%$ des gesamten Berechnungsmehraufwands bei den Use Cases ausmacht. Dies verwundert wenig, da eine einzelne Anfrage an den Chatbot bereits ca. $100.5\mathrm{g}$ CO2-Äquivalent emittiert, was ungefähr dem Vierfachen der nächsthöchsten Einzeloperation (Datenbankabfragen) und immer noch ungefähr dem Doppelten der Gigabyte-Emission des Speicheraufwands entspricht. Die Auswirkungen sind im Rahmen der verwendeten Modellierung durchaus folgerichtig.

Zu berücksichtigen ist jedoch zunächst, dass es sich hierbei um eine in Datenschutzsystemen am Ende doch eher ungewöhnliche Form des Berechnungsaufwands handelt. Außerdem ist anzumerken, dass der Chatbot aus mehreren Gründen datenschutzmäßig generell eher kritisch anzusehen ist. So wurden bei ihm am meisten Risiken von allen zulässigen Verarbeitungsvorgängen identifiziert; dem Vorgang wurde ein hohes Risiko zugeschrieben, das selbst durch die technisch-organisatorischen Maßnahmen nicht ausreichend abgemildert werden konnte, und die Zulässigkeit selber steht unter erheblichen Auflagen betreffend Freiwilligkeit. Insofern kann gut angenommen werden, dass es sich bei dem Berechnungsmehraufwand für die Anfragen an den Chatbot im Wesentlichen eher um einen *datenschutzfremden Übertrag* handelt.

**Änderungen des Nutzerverhaltens** Gegenstand der in dieser Arbeit verwendeten Modellierung von Datenschutzmaßnahmen war nicht, inwieweit diese konkrete Veränderungen in dem Verhalten betroffener Personen bei der Verwendung der Plattform induzieren. Verlangte schon die Festlegung von Annahmen zum Ausgangsverhalten einen erheblichen Argumentationsaufwand, so wäre es im Rahmen dieser Arbeit nicht möglich gewesen, zusätzlich konkrete Änderungen im Verhalten – ohne ausschließlich auf Spekulationen zurückzugreifen – herzuleiten und zu formalisieren.

Vor dem Hintergrund, dass so genannte *Dark Patterns* als sehr effektiv gelten (vgl. nur Luguri u. a. [23]), um das Verhalten von Nutzenden in eine bestimmte Richtung zu beugen, und Datenschutzmaßnahmen verlangen, dass z. B. die Entscheidung über die Nutzung des Chatbots freiwillig bleibt, dürfte anzunehmen sein, dass eine ehrliche und effektive Umsetzung der in dieser Arbeit besprochenen Maßnahmen eine erhebliche Verhaltensänderung, insbesondere die deutlich seltenere Verwendung des Chatbots zur Folge hätte. Dies dürfte aus den im vorherigen Absatz besprochenen Gründen zu einer nicht unerheblichen Verbesserung der Klimabilanz des Datenschutzes beitragen.

**Grenzen des Berechnungsmodells** Ebenfalls kein Gegenstand dieser Arbeit war es, ein allgemeines und präzises Modell zur Bestimmung der Emissionen verschiedener Operationen in einem Webserver-Kontext zusammenzustellen. Insofern wurden mangels bekannter anderweitiger Ansätze stark pauschalisierende Annahmen in Bezug auf einzeln bestimmte Operationen getroffen. Beispielsweise unterscheidet das Modell zwei Operationen gleicher Art grundsätzlich nicht voneinander. Einer Datenbank-Abfrage, die tausende Einträge löscht, wird an sich erstmal der gleiche Berechnungsaufwand und die gleiche Emission zugewiesen, wie einer Abfrage, die bloß einen einzelnen Eintrag anhand seiner ID abruft. Ähnliches gilt für HTTP-Anfragen. Diesem Ungleichgewicht wurde zwar teilweise bei der Spezifizierung der Use Cases auf begegnet (Beispiel: Use Case V), dies löst jedoch das zugrundeliegende Problem nicht.

Es ist davon auszugehen, dass die in dieser Arbeit verwendeten Faktoren in der Größenordnung plausibel sind, da sie aufgrund konkreter empirischer Erfahrungen und Beobachtungen gewählt wurden. Es wäre jedoch eine unrichtige Interpretation des Ergebnisses dieser Arbeit, beispielsweise in dieser Pauschalität zu sagen, Datenschutz würde ca. 10 Tonnen CO2-Emissionen pro Jahr verursachen.

**Plausibilität** Ungeachtet der durch den Gegenstand dieser Arbeit gebotene Aufklärungsgrenzen, einem Mangel an wissenschaftlicher Vorarbeit bei Fragen, die dieser Arbeit zugrundegelegt werden, und einer aus verschiedenen Gründen notwendigen Pauschalisierung und Typisierung dürfte realistischerweise noch davon ausgegangen werden können, dass diese Arbeit eine in der Größenordnung richtige Antwort auf die Frage nach den Klimafolgen des Datenschutzes liefert:

Technische Datenschutzmaßnahmen haben weder einen praktisch nur vernachlässigbaren Einfluss auf die Klimafolgen eines Webserver-Systems, noch haben sie überragende, schlechterdings katastrophale Klimafolgen. Ebenfalls ist davon auszugehen, dass im Rahmen dieser Arbeit aus den genannten Gründen sowohl die durch Datenschutzmaßnahmen induzierten Emissionen als auch die durch Datenschutzmaßnahmen ersparten Klimafolgen eher überschätzt wurden.

## 6.3 Bewertung und Fazit

Im Ergebnis dieser Arbeit lässt sich feststellen, dass die durch den Datenschutz induzierten Klimafolgen in dem spezifischen Szenario und mit den vorgenannten Erwägungen auch

wohl *in der Regel* nicht so hoch sind, dass dies ein Absehen von Datenschutzmaßnahmen gebieten könnte.

Insofern sind drei Punkte zu berücksichtigen:

Erstens hat diese Arbeit gezeigt, dass Datenschutz an vielen Stellen weniger durch konkrete technische Änderungen, sondern eher durch organisatorische Vor- und Maßgaben umgesetzt wird. Dies gilt nicht nur für (offensichtlich) ausschließlich organisatorische Maßnahmen, sondern auch bei einer Vielzahl von anderen Maßnahmen, die nur auf den ersten Blick technisch sind und vorrangig die richtige Benutzung oder Konfiguration der verwendeten Software zwecks Umsetzung organisatorischer Konzepte betreffen. Ein Beispiel wäre die Maßnahme 5.I. aus dem Bereich der Trennung oder die Maßnahme 8.B. aus dem Bereich der regelmäßigen Abfragen.

Zweitens ist zu beachten, dass zwischen Datenschutz und Klimaschutz kein Zielkonflikt im eigentlichen Sinne vorliegt. Beide beruhen auf dem gleichen Grundproblem, nämlich Organisationen und Strukturen, die (latent) rücksichtslos handeln (für Datenschutz vgl. Rost [28]). Beide sollen den Erhalt einer modernen, freiheitlich-demokratischen Gesellschaft gewährleisten. Und, um auf das Zitat von Thomas Hobbes im Vorwort zurückzukehren, beide dienen dem gleichen Schutzsubjekt – den Menschen.

Vor diesem Hintergrund kann weder dem Klima- noch dem Datenschutz unbedingter Vorrang vor dem jeweils anderen gewährt werden. Erstrebenswert wäre weder ein weiteres Fortschreiten der Klimakrise noch eine Gesellschaft, in der Menschen schutzlos ihren Daten und den Algorithmen ausgeliefert sind.

Drittens wird bei dem Versuch, zwischen Daten- und Klimaschutz einen zweckmäßigen Ausgleich herzustellen, zu berücksichtigen sein, dass Maßnahmen zum Klimaschutz nicht notwendigerweise eine Verschlechterung des Datenschutzniveaus bedingen müssen. Kann zum Beispiel schon der Betrieb eines Rechenzentrums an sich „klimaneutral“ erfolgen, so dürfte das gleiche für Datenschutzmaßnahmen gelten, die auf in diesem Rechenzentrum laufende Vorgänge angewendet werden. In ähnlicher Weise können auch Datenschutzmaßnahmen dem Klimaschutz dienen, wenn bspw. der Einsatz rechenintensiver Technologien (bspw. „KI“, Big Data oder „Blockchain“-Technologien) strengeren Regeln unterworfen wird.

Zu hoffen verbleibt, dass sich die Menschheit erfolgreich gegen beide Probleme behaupten kann.

# Literatur


[1] Albers und Veit. "Artikel 9 DSGVO". In: Wolff/Brink/v. Ungern-Sternberg. *BeckOK Datenschutzrecht*. 44. Ed. 1. Mai 2023.

[2] G. Apostolopoulos, V. Peris und D. Saha. "Transport layer security: how much does it really cost?" In: *IEEE INFOCOM '99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*. IEEE, 1999. DOI: 10.1109/infcom.1999.751458.

[3] BbgLDA, Hrsg. *Liste von Verarbeitungsvorgängen nach Art. 35 Abs. 4 DS-GVO*. 17. Okt. 2018. URL: https://www.lda.brandenburg.de/sixcms/media.php/9/DSFA_Muss_Liste_Allgemein_17102018.4041740.pdf (besucht am 22. 08. 2023).

[4] Katherine Calvin u. a. *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.* Techn. Ber. Juli 2023. DOI: 10.59327/ipcc/ar6-9789291691647.

[5] Ryan Chesler. *Dell PowerEdge R430*. IT Connected. 28. Juni 2017. URL: https://www.itconnected.tech/blog/dell-poweredge-r430/ (besucht am 20. 09. 2023).

[6] Vlad Coroama und Friedemann Mattern. "Zielkonflikte zwischen Umwelt- und Datenschutz. Digitalisierung nachhaltig gestalten". In: *Was Bits und Bäume verbindet*. Höfner, Anja und Frick, Vivian, Juli 2019, S. 58–60. ISBN: 978-3-96238-149-3. URL: https://www.researchgate.net/profile/Vivian-Frick/publication/34231664_Was_Bits_und_Baume_verbindet_Digitalisierung_nachhaltig_gestalten/links/5d1e1caf458515c11c12600f/Was-Bits-und-Baeume-verbindet-Digitalisierung-nachhaltig-gestalten.pdf.

[7] Datenschutzkonferenz, Hrsg. *Standard-Datenschutzmodell. Eine Methode zur Datenschutzberatung und -prüfung auf der Basis einheitlicher Gewährleistungsziele*. Version 3.0. AK Technik der Datenschutzkonferenz, 24. Nov. 2022. URL: https://www.datenschutz-mv.de/static/DS/Dateien/Datenschutzmodell/SDM-Methode-V30a.pdf (besucht am 22. 08. 2023).

[8] Dell. *PowerEdge R430 Carbon Footprint*. 2019. URL: https://i.dell.com/sites/csdocuments/CorpComm_Docs/en/carbon-footprint-poweredge-r430.pdf (besucht am 20. 09. 2023).

[9] Yannik Dittmar u. a. *HPI Data Center Climate Footprint*. 2023. URL: https://climateboard-bptr1.hpi.de/.

[10] Yannik Dittmar u. a. *HPI Data Center Climate Footprint - Operational Carbon*. 2023. URL: https://climateboard-bptr1.hpi.de/documentation/books/hpi-data-center-climate-footprint/page/operational-carbon.

[11] Charlotte Freitag u. a. "The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations". In: *Patterns* 2.9 (Sep. 2021), S. 100340. DOI: 10.1016/j.patter.2021.100340.

[12] Fujitsu. *Data Sheet - Fujitsu PRIMERGY RX2530 M5 Rack Server*. 2. Sep. 2023. URL: https://sp.ts.fujitsu.com/dmsp/Publications/public/ds-py-rx2530-m5.pdf (besucht am 20. 09. 2023).

[13] Arthur Goldberg, Robert Buff und Andrew Schmitt. "A Comparison of HTTP and HTTPS Performance". In: (). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1cd89de5cf0e618924c73ac2b060104b7076f0b7.

[14] Thaier Hayajneh u. a. "Performance and Information Security Evaluation with Firewalls". In: *International Journal of Security and Its Applications* 7.6 (Nov. 2013), S. 355–372. DOI: 10.14257/ijsia.2013.7.6.36.

[15] Garth Heward u. a. "Assessing the Performance Impact of Service Monitoring". In: *2010 21st Australian Software Engineering Conference*. IEEE, 2010. DOI: 10.1109/aswec.2010.28.

[16] HPE Compute. *HPE product carbon footprint - HPE ProLiant DL380 Gen10 Server*. 2023. URL: https://www.hpe.com/psnow/doc/a50004545enw (besucht am 20. 09. 2023).

[17] ISO, Hrsg. *ISO/IEC 9899:2018. Information technology — Programming languages — C*. Juni 2018.

[18] Kassenärztliche Bundesvereinigung. *GESUNDHEITSDATEN - Die meisten gehen 3 bis 5 Mal pro Jahr zum Arzt*. 2021. URL: https://gesundheitsdaten.kbv.de/cms/html/24044.php (besucht am 19. 09. 2023).

[19] Kassenärztliche Bundesvereinigung. *GESUNDHEITSDATEN - Die Wartezeit ist für die meisten kurz*. 2021. URL: https://gesundheitsdaten.kbv.de/cms/html/24045.php (besucht am 19. 09. 2023).

[20] Kassenärztliche Bundesvereinigung. *GESUNDHEITSDATEN - Mehr als 548.000 Ärzte in Deutschland*. 2021. URL: https://gesundheitsdaten.kbv.de/cms/html/17077.php (besucht am 19. 09. 2023).

[21] Stefan Krempl. *Unfallmediziner: DSGVO-Auslegung "gefährdet Menschenleben"*. 28. Okt. 2021. URL: https://www.heise.de/news/Unfallmediziner-DSGVO-Auslegung-gefaehrdet-Menschenleben-6233776.html.

[22] Percy Liang u. a. "Holistic Evaluation of Language Models". In: *Transactions on Machine Learning Research (TMLR), 2023* (16. Nov. 2022). arXiv: 2211.09110v2 [cs.CL].

[23] Jamie Luguri und Lior Jacob Strahilevitz. "Shining a Light on Dark Patterns". In: *Journal of Legal Analysis* 13.1 (Jan. 2021), S. 43–109. DOI: 10.1093/jla/laaa006.

[24] Materials System Laboratory. *PAIA Research Approach*. URL: https://msl.mit.edu/projects/paia/paia-research-approach (besucht am 14. 10. 2023).

[25] Dmitrij Melkov, Arunas Saltis und Sarunas Paulikas. "Performance Testing of Linux Firewalls". In: *2020 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*. IEEE, Apr. 2020. DOI: 10.1109/estream50540.2020.9108868.

[26] Britta Alexandra Mester. "Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten". In: Taeger/Gabel. *DSGVO - BDSG - TTDSG*. 4. Aufl. 2022.

[27] Elsa Olivetti und Randolph Kirchain. “A Product Attribute to Impact Algorithm to Streamline IT Carbon Footprinting”. In: *Design for Innovative Value Towards a Sustainable Society*. Springer Netherlands, 2012, S. 747–749. DOI: 10.1007/978-94-007-3010-6_151.

[28] Martin Rost. *Das Standard-Datenschutzmodell (SDM). Einführung, Hintergründe und Kontexte zum Erreichen der Gewährleistungsziele*. Springer Fachmedien Wiesbaden, 2022. ISBN: 978-3-658-38879-9. DOI: 10.1007/978-3-658-38880-5.

[29] Schulz. “Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten”. In: Gola/Heckmann. *Datenschutz-Grundverordnung - Bundesdatenschutzgesetz*. 3. Auflage. 2022.

[30] Christian Spöcker. *Intensivmediziner: Datenschutz hat optimale Bekämpfung der Corona-Pandemie verhindert*. 30. Dez. 2021. URL: http://web.archive.org/web/20220825121103/https://www.swr.de/swraktuell/radio/intensivmediziner-datenschutz-hat-optimale-bekaempfung-der-corona-pandemie-verhindert-100.html.

[31] Statistisches Bundesamt. *Bevölkerung nach Nationalität und Geschlecht (Quartalszahlen)*. 2023. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/liste-zensus-geschlecht-staatsangehoerigkeit.html (besucht am 19. 09. 2023).

[32] Paul Strobach. *Datenschutz-Philosophie*. 2023. URL: https://luap42.de/notes/privacy-philosophy/ (besucht am 22. 08. 2023).

[33] “Summary for Policymakers”. In: *Climate Change 2022 – Impacts, Adaptation and Vulnerability*. Cambridge University Press, Juni 2023, S. 3–34. DOI: 10.1017/9781009325844.001.

[34] Hugo Touvron u. a. *Llama 2: Open Foundation and Fine-Tuned Chat Models*. 18. Juli 2023. arXiv: 2307.09288v2 [cs.CL].

[35] Hugo Touvron u. a. *LLaMA: Open and Efficient Foundation Language Models*. 27. Feb. 2023. arXiv: 2302.13971v1 [cs.CL].

[36] Truefoundry. *Benchmarking Llama-2-7B*. 12. Sep. 2023. URL: https://blog.truefoundry.com/llama-2-benchmarks/.

[37] Hendrik Wieduwilt. *Datenschutz frisst Nikolaus*. 29. Okt. 2021. URL: https://www.n-tv.de/politik/Datenschutz-frisst-Nikolaus-article22895809.html.

[38] Alexander Wulfers. *Wann hört der Ärger mit den Cookies endlich auf?* 24. Jan. 2023. URL: https://www.faz.net/aktuell/wirtschaft/cookie-banner-wann-hoert-der-aerger-endlich-auf-18617577.html.

# Anhang A

# Datenschutz-Folgenabschätzung

Auf den folgenden Seiten ist die im Rahmen dieser Bachelorarbeit vorgenommene Datenschutz-Folgenabschätzung mitsamt der Anlage 1 – Liste aller technisch-organisatorischen Maßnahmen – aufgeführt.

# Datenschutz-Folgenabschätzung

**Beschreibung des Verarbeitungssystems**

Dieser DSFA liegt das folgende (hypothetische) Szenario aus meiner Bachelorarbeit zugrunde:

> *Ein Unternehmen möchte eine Plattform anbieten, über die Patient*innen so einfach wie möglich Arzttermine buchen können, und zwar am Besten bei allen Praxen. Da das Unternehmen eine Ausgründung des HPI ist, soll die Plattform auf dem HPI-Rechenzentrum laufen. Patient*innen können auf der Plattform Termine bei kooperierenden, auf der Plattform eingelisteten Ärzt*innen buchen und gebuchte Termine ändern (Aufhebung, Verlegung). Sie werden an den Termin rechtzeitig erinnert und erhalten über die Plattform einen Buchungscode, mit dem sie sich bei der Anmeldung ausweisen können.*
>
> *In einem nächsten Schritt möchte das Unternehmen die Plattform durch „KI“ anreichern. Damit sollen unnötige Arzttermine z. B. bei einer leichten Erkältung vermieden werden. Hierfür soll vor die Terminbuchung ein Chatbot geschaltet werden, der den Patient*innen einige Fragen zu ihrem Terminsgrund, bzw. ggf. zu ihren medizinischen Problemen stellt und dann entweder -- in einfachen und unproblematischen Fällen wie der besagten leichten Erkältung -- auf seine Diagnose und übliche alltägliche Behandlungsmittel hinweist, oder in allen anderen Fällen automatisch Ärzt*innen vorschlägt, bei denen man unmittelbar auf der Plattform einen Termin buchen sollte.*

Hieraus ergeben sich diese, im Folgenden einzeln betrachteten, Verarbeitungsvorgänge:

1. Anzeige von Arztinformationen
2. Buchung eines Termins
3. Verwaltung eines Termins
4. Terminerinnerung
5. Vor-Ort-Identifizierung
6. Erzeugung von Trainingsdaten
7. Training
8. Konsultation

**Notwendigkeit einer DSFA**

Eine DSFA ist notwendig nach Art. 35 DSGVO, wenn die Form des Verarbeitungssystems voraussichtlich ein hohes Risiko für die betroffenen Personen zur Folge hat.

Es liegt bereits der Regelfall des Art. 35 Abs. 3 Buchst. b) DSGVO vor, nämlich die umfangreiche Verarbeitung besonderer Kategorien von personenbezogenen Daten gemäß Art. 9 Abs. 1 DSGVO (hier: Gesundheitsdaten).

Zudem sind auch mehrere Schwellwertkriterien nach WP248 erfüllt, nämlich Nr. 4 – vertrauliche Daten oder höchst persönliche Daten –, Nr. 5 – Datenverarbeitung in großem Umfang –, sowie bzgl. einzelner Systembestandteile auch die Nr. 2 – automatisierte Entscheidungsfindung mit Rechtswirkung oder ähnlich bedeutsamer Wirkung –, nämlich insoweit das System Einfluss auf die Durchführung oder Nichtdurchführung medizinischer Maßnahmen ausübt, und Nr. 8 – innovative Nutzung oder Anwendung neuer technologischer oder organisatorischer Lösungen –, nämlich bzgl. der „KI“-Technologie.

Daher hat das gesamte Verarbeitungssystem ein hohes Risiko für die Rechte und Freiheiten betroffener Personen, sodass eine DSFA insgesamt aus rechtlichen Gründen notwendig ist.

Eine DSFA ist zudem, ohne dass es darauf noch ankäme, auch für die Bearbeitung des Szenarios in der Bachelorarbeit zweckmäßig und geboten.

**Legende**

Im Rahmen dieser DSFA findet die folgende Adressierung statt:

| | |
|---|---|
| 1, 2, 3, … | Verarbeitungsvorgänge |
| 1.a, 1.b, 1.c, … | Identifizierte Risiken eines Verarbeitungsvorgangs |
| 1.a.1, 1.a.2, 1.a.3, … | Generische TOM (nach SDM) für das jeweilige Risiko |
| 1.A, 1.B, 1.C, … | Spezifische TOM für einen Verarbeitungsvorgang |

**Verarbeitungsvorgänge**

### 1. Anzeige von Arztinformationen (Termine)

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 1. | Anzeige von Arztinformationen | In der Anwendung werden Ärzt*innen in durchsuchbarer Form, insbesondere nach Fachbereich und Ort, angezeigt; für jeden Eintrag werden u. a. Spezialisierung, Adresse und Kontaktmöglichkeiten angezeigt. | Vermittlung von Patient*innen |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Verarbeitungsvorgang 1.1 kann auf die Erforderlichkeit für die Erfüllung eines Vertrag (Art. 6 Abs. 1 Buchst. b) DSGVO) zwischen dem Anbieter der Plattform und den jeweiligen medizinischen Fachpersonen gestützt werden, welcher die Vermittlung von Patient*innen betrifft.

*Bewertung der Notwendigkeit und Angemessenheit*

Der Verarbeitungsvorgang 1 ist für seine Zwecke notwendig und angemessen. Für die Erfüllung des Zwecks der Vermittlung von Patient*innen gibt es keine wenigstens genauso effektive aber mildere Verarbeitungsform. Eine „blinde“ Zuweisung anhand bestimmter Angaben der Patient*innen ist weniger effektiv, da sie keine Patient*innenbindung zulässt und fehlerhaft sein kann. Andere, potentiell mildere Verarbeitungsformen sind nicht ersichtlich. Zudem ist die Verarbeitung auch angemessen, hier ist zu berücksichtigen, dass die Eintragung in das Register freiwillig aufgrund einer Vereinbarung erfolgt und die Interessen der betroffenen Person mit den Zwecken der Verarbeitung gleichgerichtet sind.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Es besteht das Risiko, dass die Intervenierbarkeit und die Integrität nicht vollständig gewährleistet werden, indem die bereitgestellten Daten unrichtig sind und nicht korrigiert werden (können). Dies führt mittelbar zu wirtschaftlichen und rechtlichen Nachteilen für die betroffenen Ärzt*innen, da ggf. keine Termine gebucht werden können, gebuchte Termine nicht wahrgenommen werden können, oder gegen berufsrechtliche Vorschriften zur Werbung verstoßen werden könnte.

(b) Unter Verletzung der Gewährleistungsziele der Integrität und der Transparenz könnte die Darstellung verzerrt sein, was auch zu den ebengenannten mittelbaren Nachteilen führen könnte.

(c) Unter Verletzung der Ziele der Datenminimierung, der Nichtverkettung und der Intervenierbarkeit könnte es durch eine überlange oder übermäßige Speicherung/Veröffentlichung von Angaben zu einer illegitimen Verkettung kommen, z. B. durch die weitere Veröffentlichung von Informationen nach einer Praxisaufgabe.

(d) Unter Verletzung des Grundsatzes der Datenminimierung könnten unnötige Informationen verarbeitet werden, wodurch in unberechtigter Weise in die Rechte und Freiheiten der betroffenen Person eingegriffen werden würde.

Eine Schwellwertanalyse der Risiken ergibt, dass ausschließlich das Kriterium „Datenverarbeitung in großem Umfang“ erfüllt ist. Zudem betreffen die Daten ausschließlich die Sozialsphäre, sodass das Risiko für Verarbeitungsvorgang 1 insgesamt als **normal** einzustufen ist.

*Vorgeschlagene Abhilfemaßnahmen*

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Die Risiken eines Verstoßes gegen die Ziele der Intervenierbarkeit und Integrität können vorliegend geeignet durch technisch-organisatorische Maßnahmen wie eine Möglichkeit zum Ändern von eingetragenen Daten, (2.) Prozesse zur Aufrechterhaltung der Aktualität von Daten und (3.) das Löschen oder Berichtigen falscher Daten gemildert werden.

(b) Die Risiken einer Verletzung gegen die Ziele der Integrität und Transparenz können vorliegend geeignet durch (1.) Prozesse zur Aufrechterhaltung der Aktualität von Daten, (2.) das Löschen oder Berichtigen falscher Daten, (3.) die Dokumentation der Faktoren, die für eine Profilierung, zum Scoring oder für teilautomatisierte Entscheidungen genutzt werden, sowie (4.) die Dokumentation der Bestandteile von Verarbeitungstätigkeiten insbesondere der Geschäftsprozesse, Datenbestände, Datenflüsse und Netzpläne, dafür genutzte IT-Systeme, Betriebsabläufe, Beschreibungen von Verarbeitungstätigkeiten, Zusammenspiel mit anderen Verarbeitungstätigkeiten gemildert werden.

(c) Das Risiko einer Verletzung der Ziele der Datenminimierung und der Nichtverkettung kann geeignet durch (1.) die Festlegung und Umsetzung eines Löschkonzepts sowie (2.) durch die Implementierung standardisierter Abfrage- und Dialogschnittstellen für Betroffene zur Geltendmachung und/oder Durchsetzung von Ansprüchen abgemildert werden.

(d) Dem Risiko eines Verstoßes gegen das Ziel der Datenminimierung könnte vorliegend geeignet (1.) durch die Reduzierung von erfassten Attributen der betroffenen Personen, sowie (2.) durch die Festlegung von Voreinstellungen für betroffene Personen, die die Ver-

arbeitung ihrer Daten auf das für den Verarbeitungszweck erforderliche Maß beschränken, begegnet werden.

Ergebnisse:

Unter zweckmäßiger Kombination der bei 1.a bis 1.d aufgeführten Einzelmaßnahmen bieten sich für den Verarbeitungsvorgang 1 insgesamt die folgenden Maßnahmen an:

A) Implementierung einer Möglichkeit zum Ändern oder Löschen von eingetragenen Daten (1.a.1., 1.c.2.)

B) Festlegung und Umsetzung eines Konzepts mit Prozessen zur Löschung oder Berichtigung falscher oder veralteter Daten (1.a.2., 1.a.3, 1.b.1., 1.b.2., 1.c.1.)

C) Dokumentation der Faktoren und Bestandteile der Darstellung von Ärzteprofilen (1.b.3., 1.b.4.)

D) Beschränkung der (standardmäßig) erfassten Attributen bzgl. Ärzt*innen und Arztpraxen auf den notwendigen Inhalt (1.d.1., 1.d.2.)

**2. Buchung eines Termins (Termine)**

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 2. | Buchung eines Termins | Patient*innen können in der Anwendung bei einer Praxis einen Termin buchen. Die Praxis wird über den Termin informiert. Bei der Terminbuchung können bereits einfache Angaben über das medizinische Problem (z. B. Hautausschlag) oder das Terminziel (z. B. Impfung gegen COVID19) gemacht werden. | Vermittlung von Patient*innen<br><br>Geschäftsanbahnung |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Rechtsgrundlagen für den Verarbeitungsvorgang 2 ist die Erforderlichkeit für die Erfüllung bzw. die Vorbereitung eines Behandlungsvertrags mit den jeweiligen betroffenen Personen. Dies ergibt sich daraus, dass medizinische Dienstleistungen üblicherweise aufgrund eines Termins angeboten werden und über den Termin Einigkeit hergestellt werden muss, sodass die Buchung eines Termins (und ggf. dessen Änderung) als notwendige Vorbedingung für die Erfüllung des Vertrags angesehen werden kann.

Der Verarbeitungsvorgang 2 betrifft jedoch auch Gesundheitsdaten im Sinne des Art. 9 Abs. 1 DSGVO, sodass ein Ausnahmegrund nach Art. 9 Abs. 2 DSGVO bestehen muss. Bereits die Buchung eines Arzttermins kann in erheblicher Weise stigmatisierend wirken, man denke an einen Termin bei einer Ärztin oder einem Arzt für Geschlechtskrankheiten oder bzgl. der Planung oder Vornahme eines Schwangerschaftsabbruchs. Auch die Information, dass ein entsprechender Termin gebucht wird, beinhaltet schon das spezifische Risiko von Gesundheitsdaten, sodass nicht erst eine bestätigte Diagnose oder eine begonnene Therapie unter den erhöhten Schutz des Art. 9 DSGVO fallen können.

Hier kommt im Ergebnis nur Art. 9 Abs. 2 Buchst. h) DSGVO, also die Erforderlichkeit für eine vertraglich vereinbarte Diagnose, Vorsorge oder Behandlung im Gesundheitsbereich in Betracht. Eine den Anforderungen der Art. 7, 9 Abs. 2 Buchst. a) DSGVO entsprechende Einwilligung dürfte bei einem solchen Massengeschäft, wie es hier vorliegt, kaum rechtswirksam einzuholen sein. Andere Ausnahmegründe sind ersichtlich nicht einschlägig. Aus den oben bereits für Art. 6 Abs. 1 Buchst. b) DSGVO angeführten Gründen dürfte für für den Vorgang 2 auch Art. 9 Abs. 2 Buchst. h) DSGVO geeignet sein.

Hieraus ergibt sich jedoch eine Besonderheit im weiteren Verfahren. Art. 9 Abs. 2 Buchst. h) DSGVO gilt nur unter dem Vorbehalt, dass die Verarbeitung unter der Verantwortung von medizinischem Fachpersonal erfolgt, welches einer gesetzlichen Geheimhaltungs-

pflicht unterliegt. Auf einen solchen Status wird sich das Unternehmen nicht berufen können, weshalb das Vertragsverhältnis (s. zuvor bei 1.) so ausgestaltet werden muss, dass das jeweilige medizinische Fachpersonal für die Verarbeitungsvorgänge in rechtlicher und tatsächlicher Hinsicht ausschließlich verantwortlich ist, d. h. die Zwecke und Mittel der Verarbeitung (umfassend) bestimmt. Das Unternehmen kann seine Dienstleistung nur im Wege der Auftragsverarbeitung anbieten.

*Bewertung der Notwendigkeit und Angemessenheit*

Der Verarbeitungsvorgang 1.2 ist für seine Zwecke notwendig, denn, wie bereits zuvor beschrieben, stellt ein Termin gerade die übliche Gegebenheit dar, bei der medizinische Dienstleistungen erbracht werden. Der Erfolg der Vermittlung von Patient*innen bzw. der Anbahnung einer Geschäftsbeziehung bemisst sich daher gerade mit der Buchung von Terminen. Das mildere Mittel, nur Kontaktmöglichkeiten des medizinischen Portals für eine manuelle Geschäftsanbahnung bereitzustellen, ist ersichtlich weniger geeignet, den Erfolg herbeizuführen, da dies für die Patient*innen unbequemer wäre, sodass sie mit höherer Wahrscheinlichkeit keinen Termin buchen würden. Die Angemessenheit ist grundsätzlich auch zu bejahen, da die Verarbeitung im Interesse und auf Anfrage der betroffenen Person erfolgt, da diese ja explizit die Terminbuchung anstößt und – wie man aus ihrem Verhalten schließen kann – wohl gerne einen Termin hätte. Zu beachten ist jedoch, dass durch die Verbreitung der Plattform ein Nutzungszwang entstehen könnte, der der Freiwilligkeit der Nutzung der Plattform entgegensteht. Daher dürfte es der Grundsatz der Angemessenheit gebieten, dass die Nutzung der Plattform stets auch tatsächlich freiwillig erfolgt, d. h. plattformfreie Möglichkeiten der Terminbuchung bestehen.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Unter Verletzung des Gewährleistungsziels der Vertraulichkeit und der Nichtverkettung könnten hochsensitive Gesundheitsdaten Unbefugten bekannt und von diesen ggf. weiterverarbeitet werden.

(b) Zudem besteht das Risiko, dass unter Verletzung des Ziels der Verfügbarkeit die Terminbuchung oder der Termin nicht oder nicht mehr verfügbar ist, weshalb eine dringende medizinische Maßnahme oder ein Notfall nicht oder nicht, wie geplant, durchgeführt werden kann.

(c) Unter Verletzung des Grundsatzes der Datenminimierung könnten vor dem Termin unnötig viele Informationen abgefragt werden, wodurch in unberechtigter Weise in die Rechte und Freiheiten der betroffenen Person eingegriffen werden würde.

(d) Unter Verletzung des Ziels der Transparenz könnte nicht eindeutig zu erkennen sein, wer welche Daten einsehen kann, sodass nicht ausgeschlossen werden kann, dass sich das Risiko a verwirklicht.

(e) Ein weiteres Risiko besteht zudem darin, dass unter Verletzung des Gewährleistungsziels der Intervenierbarkeit ein Termin nicht oder nur schwer aufgehoben oder geändert werden könnte, wenn dies aus tatsächlichen Umständen erforderlich wird.

Eine Schwellwertanalyse der Risiken ergibt, dass – wie oben bereits begründet – „vertrauliche oder höchstsensitive Daten“ betroffen sind. Zudem erfolgt wohl auch die „Datenverarbeitung in größerem Umfang“ – sowohl zeitlich als auch örtlich –. Soweit Terminbuchungen minderjähriger Patient*innen betroffen sind, liegen außerdem „Daten zu schutzwürdigen Betroffenen“ vor. Zu berücksichtigen ist auch, dass ein Teil der Daten die Intimsphäre betrifft, weshalb das Risiko als **hoch** einzustufen ist.

*Vorgeschlagene Abhilfemaßnahmen*

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Dem Risiko eines Verstoßes gegen die Ziele der Vertraulichkeit und Nichtverkettung könnte geeignet (1.) durch die Festlegung eines Berechtigungs- und Rollenkonzeptes nach dem Erforderlichkeitsprinzip auf der Basis eines Identitätsmanagements durch die verantwortliche Stelle, (2.) durch die Implementierung eines sicheren Authentifizierungsverfahrens, (3.) durch die Eingrenzung der zulässigen Personalkräfte auf solche, die nachprüfbar zuständig (örtlich, fachlich), fachlich befähigt, zuverlässig (ggf. sicherheitsüberprüft) und formal zugelassen sind sowie keine Interessenskonflikte bei der Ausübung aufweisen, (4.) durch spezifizierte, für die Verarbeitungstätigkeit ausgestattete Umgebungen (Gebäude, Räume), (5.) durch die Verschlüsselung von gespeicherten oder transferierten Daten sowie Prozesse zur Verwaltung und zum Schutz der kryptografischen Informationen (Kryptokonzept), (6.) durch den Schutz vor äußeren Einflüssen (Spionage, Hacking), (7.) durch die Einschränkung von Verarbeitungs-, Nutzungs- und Übermittlungsrechten, (8.) durch die Trennung nach Organisations-/Abteilungsgrenzen, sowie (9.) durch die Trennung mittels Rollenkonzepten mit abgestuften Zugriffsrechten auf der Basis eines Identitätsmanagements durch die verantwortliche Stelle und eines sicheren Authentifizierungsverfahrens begegnet werden.

(b) Das Risiko für die Verfügbarkeit kann vorliegend geeignet (1.) durch den Schutz vor äußeren Einflüssen (Schadsoftware, Sabotage, höhere Gewalt), (2.) durch die Redundanz von Hard- und Software sowie Infrastruktur, sowie (3.) durch die Erstellung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit abgemildert werden.

(c) Dem Risiko für die Datenminimierung kann geeignet (1.) durch die Reduzierung von erfassten Attributen der betroffenen Personen sowie (2.) durch die Festlegung von Voreinstellungen für betroffene Personen, die die Verarbeitung ihrer Daten auf das für den Verarbeitungszweck erforderliche Maß beschränken, begegnet werden.

(d) Dem Risiko für die Transparenz kann durch (1.) die Dokumentation der Bestandteile von Verarbeitungstätigkeiten insbesondere der Geschäftsprozesse, Datenbestände, Datenflüsse und Netzpläne, dafür genutzte IT-Systeme, Betriebsabläufe, Beschreibungen

von Verarbeitungstätigkeiten, Zusammenspiel mit anderen Verarbeitungstätigkeiten, (2.) die Protokollierung von Zugriffen und Änderungen, sowie (3.) die Benachrichtigung von Betroffenen bei Datenpannen oder bei Weiterverarbeitungen zu einem anderen Zweck angemessen begegnet werden.

(e) Dem Risiko für die Intervenierbarkeit kann durch (1.) die Identifizierung und Authentifizierung der Personen, die Betroffenenrechte wahrnehmen möchten, sowie (2.) die Möglichkeit der Änderung und Löschung von eingetragenen Daten angemessen begegnet werden.

Ergebnis:

Unter zweckmäßiger Kombination der bei 2.a bis 2.e aufgeführten Einzelmaßnahmen bieten sich für den Verarbeitungsvorgang 2 insgesamt die folgenden Maßnahmen an:

A) Beschränkung des Zugriffs zu den Terminen und ihre Verarbeitung auf authentifizierte und (anhand ihrer Prozessrolle sowie im Einzelfall) berechtigte Personen (2.a.1., 2.a.2., 2.a.7., 2.a.9., 2.e.1.)

B) Einführung von Anforderungen an Personen (außer den Patient*innen), die auf Termineinträge zugreifen wollen, wie bspw. nachprüfbare Zuständigkeit, fachliche Befähigung, Zuverlässigkeit oder die förmliche Verpflichtung zur Verschwiegenheit (2.a.3.)

C) Beschränkung des (technischen wie physischen) Zugangs und Schutz des Rechenzentrums, in dem die Verarbeitungstätigkeit stattfindet, wie z. B. durch eine Firewall, ein Schadsoftware-Monitoring oder eine Sicherheitsschleuse (2.a.2., 2.a.4., 2.a.6., 2.b.1.)

D) Einsatz von Transportverschlüsselung bei Anfragen an die Plattform, Verschlüsselung der gespeicherten Daten, sowie Einführung eines Kryptokonzepts (2.a.5.)

E) Logische und ggf. physische Trennung der Daten nach Organisations-/Abteilungsgrenzen (2.a.4., 2.a.8., 2.a.9.)

F) Herstellung von Redundanz von Hard- und Software sowie Infrastruktur sowie Festlegung und regelmäßige Kontrolle eines Verfahrens zur Wiederherstellung der Verarbeitungstätigkeit (2.b.2., 2.b.3.)

G) Reduktion der standardmäßig erfassten Angaben bei der Terminbuchung auf notwendige Angaben wie Zeitpunkt und ausgewählte Praxis (2.c.1., 2.c.2.)

H) Protokollierung von Zugriffen und Änderungen sowie Benachrichtigung bei unberechtigten Maßnahmen (2.d.2., 2.d.3.)

I) Dokumentation der Möglichkeiten, wie auf die Termindaten Zugriff erhalten werden kann, welche Schutzmaßnahmen dagegen bestehen und wie ein (berechtigter oder unberechtigter) Zugriff nachvollzogen werden kann (2.d.1.)

J) Implementierung einer Möglichkeit der Änderung und Löschung von eingetragenen Daten durch die betroffenen Personen (2.e.2.)

Da für die Verarbeitungstätigkeit ein hohes Risiko festgestellt worden ist, müssen diese Maßnahmen besonders wirksam umgesetzt werden, z. B. muss bei 2.D eine dem Stand der Technik entsprechende und nach den gegenwärtigen Erkenntnissen auch auf absehbare Zeit sichere Verschlüsselungstechnik verwendet werden oder bei 2.H eine revisionssichere Protokollierung erfolgen.

### 3. Verwaltung eines Termins (Termine)

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 3. | Verwaltung eines Termins | Gebuchte Termine können von den Patient*innen angesehen, geändert und aufgehoben werden. | Flexibilität für Patient*innen |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Es gilt umfassend das bereits bei Verarbeitungsvorgang 2 gesagte. Die Herstellung der Einigkeit über den Termin ist wesentliche Voraussetzung für die Durchführung eines Behandlungsvertrags. Wenn sich hier aufgrund der Umstände zwingende Änderungen ergeben, dann müssen diese auch berücksichtigt werden können, weshalb der Vorgang nach Art. 6 Abs. 1 Buchst. b) DSGVO gerechtfertigt ist.

Aus den dort genannten Gründen liegen auch hier Gesundheitsdaten vor, jedoch ist – aus den gleichen Gründen wie zuvor – auch hier der Ausnahmegrund des Art. 9 Abs. 2 Buchst. h) DSGVO anwendbar.

*Bewertung der Notwendigkeit und Angemessenheit*

Verarbeitungsvorgang 3 ist für seine Zwecke notwendig und angemessen. Der Verarbetungsvorgang stellt eine Ausprägung des Gewährleistungsziels der Intervenierbarkeit für den Verarbeitungsvorgang 2 dar und dient der Flexibilität der Patient*innen. Mildere Verarbeitungsvorgänge sind nicht ersichtlich. Die Verarbeitung ist auch angemessen, da das Interesse der betroffenen Personen mit dem Zweck gleichgerichtet ist; die Verarbeitung ist ausschließlich positiv für die Patient*innen.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Es besteht das Risiko, dass unter Verletzung des Gewährleistungsziels der Integrität unberechtigte Personen Zugriff erhalten und dadurch einen Termin ohne Wissen und/oder gegen den Willen der betroffenen Person ändern oder aufheben.

(b) Zudem besteht wie bei 2.a das Risiko, dass die Termine von unberechtigten Personen eingesehen werden können.

(c) Ein weiteres Risiko des Verarbeitungsvorgangs ist, dass unter Verletzung des Gewährleistungsziels der Transparenz nicht nachvollziehbar ist, von wem wann welche Daten verändert wurden, sodass nicht ausgeschlossen werden kann, dass sich das Risiko a verwirklicht.

Eine Schwellwertanalyse führt zu dem Ergebnis, dass „vertrauliche oder höchstpersönliche Daten“ betroffen sind und dass es sich wohl um eine „Datenverarbeitung in größerem Umfang“ handelt. Zudem kann durch eine unbeabsichtigte oder unberechtigte Stornierung eines Termins die Patientin bzw. der Patient „an der Ausübung eines Rechts oder der Nutzung einer Dienstleistung bzw. Durchführung eines Vertrags gehindert“ werden. Zu berücksichtigen ist schließlich noch, dass ein Teil der Daten die Intimsphäre betrifft, weshalb das Risiko insgesamt als **hoch** einzustufen ist.

*Vorgeschlagene Abhilfemaßnahmen*

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Dem Risiko für die Integrität kann vorliegend auf geeignete Weise (1.) durch die Einschränkung von Schreib- und Änderungsrechten, (2.) durch eine dokumentierte Zuweisung von Berechtigungen und Rollen, (3.) durch Prozesse zur Identifizierung und Authentifizierung von Personen und Gerätschaften, sowie (4.) durch den Schutz vor äußeren Einflüssen (Spionage, Hacking) begegnet werden.

(b) Dem Risiko für die Vertraulichkeit und Nichtverkettung kann durch die für 2.a vorgesehenen Maßnahmen begegnet werden.

(c) Dem Risiko für die Transparenz kann (1.) durch die Protokollierung von Zugriffen und Änderungen, (2.) durch Versionierung, sowie (3.) durch eine Dokumentation der Verarbeitungsprozesse mittels Protokollen auf der Basis eines Protokollierungs- und Auswertungskonzepts angemessen begegnet werden.

Ergebnis:

Unter zweckmäßiger Kombination der bei 3.a bis 3.e aufgeführten Einzelmaßnahmen bieten sich für den Verarbeitungsvorgang 3 insgesamt die folgenden Maßnahmen an:

A) die Maßnahmen 2.A, 2.B, 2.C, 2.D und 2.E:

   A) Beschränkung des Zugriffs zu den Terminen und ihre Verarbeitung auf authentifizierte und (anhand ihrer Prozessrolle sowie im Einzelfall) berechtigte Personen (3.b.*., 3.a.1., 3.a.3.)

   B) Einführung von Anforderungen an Personen (außer den Patient*innen), die auf Termineinträge zugreifen wollen, wie bspw. nachprüfbare Zuständigkeit, fachliche Befähigung, Zuverlässigkeit oder die förmliche Verpflichtung zur Verschwiegenheit (3.b.*.)

   C) Beschränkung des (technischen wie physischen) Zugangs und Schutz des Rechenzentrums, in dem die Verarbeitungstätigkeit stattfindet, wie z. B. durch eine Firewall, ein Schadsoftware-Monitoring oder eine Sicherheitsschleuse (3.b.*., 3.a.4.)

   D) Einsatz von Transportverschlüsselung bei Anfragen an die Plattform, Verschlüsselung der gespeicherten Daten, sowie Einführung eines Kryptokonzepts (3.b.*.)

B) Logische und ggf. physische Trennung der Daten nach Organisations-/Abteilungsgrenzen (3.b.*.)

C) Dokumentation der Zuweisung von Berechtigungen und Rollen (3.a.2.)

D) Entwicklung, Dokumentation und Umsetzung eines Konzepts zur Protokollierung von Zugriffen und Änderung inkl. einer (wiederherstellbaren) Versionierung der Daten (3.c.1., 3.c.2., 3.c.3.)

Da für die Verarbeitungstätigkeit ein hohes Risiko festgestellt worden ist, müssen diese Maßnahmen besonders wirksam umgesetzt werden, z. B. muss bei 3.A.D eine dem Stand der Technik entsprechende und nach den gegenwärtigen Erkenntnissen auch auf absehbare Zeit sichere Verschlüsselungstechnik verwendet werden oder bei 2.C eine revisionssichere Protokollierung erfolgen.

**4. Terminerinnerung (Termine)**

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 4. | Terminerinnerung | Die Patient*innen werden zwei Tage vor dem Termin an diesen mit einer E-Mail erinnert. Die E-Mail wird an die im Konto gespeicherte E-Mail-Adresse gesendet. | (Sicherung der) Geschäftsanbahnung |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Für den Verarbeitungsvorgang 4 kommt die Rechtsgrundlage des Art 6 Abs. 1 Buchst. b) DSGVO nicht in Betracht, denn die Verarbeitung ist objektiv nicht für die Erfüllung des Behandlungsvertrags erforderlich; es ist den Beteiligten mit zumutbarem Aufwand möglich, ihre Vertragspflichten ohne eine solche Erinnerung zu erfüllen.

In Betracht kommen jedoch die Rechtsgrundlagen der Einwilligung (Art. 6 Abs. 1 Buchst. a) DSGVO) sowie des berechtigten Interesses, hier das wirtschaftliche Interesse der Ärzt*innen an der tatsächlichen Wahrnehmung des Termins (Art. 6 Abs. 1 Buchst. f) DSGVO). Dies kommt wegen Art. 7 Abs. 3 und Art. 21 Abs. 1 DSGVO im Ergebnis der Unterscheidung zwischen einem Opt-In und einem Opt-Out-System gleich. Nach Art. 25 DSGVO sollen jedoch Verarbeitungsvorgänge so gestaltet werden, dass datenschutzfreundliche Voreinstellungen gewählt werden. Zudem handelt es sich bei der Erinnerung um eine klassische „Zusatzleistung“, sodass einer Einwilligung höchstens geringfügige Bedenken im Hinblick auf potentielle Machtasymmetrien entgegenstehen dürften. Daher erscheint eine ausdrückliche Einwilligung hier als Rechtsgrundlage vorzugswürdiger.

Daraus folgt dann auch der Ausnahmegrund, nämlich die Einwilligung nach Art. 9 Abs. 2 Buchst. a) DSGVO. Die für die anderen Vorgänge angeführte Ausnahme kommt hier wiederum mangels objektiver Erforderlichkeit nicht in Betracht. Eine Einwilligung erscheint aus den vorab genannten Gründen nicht unangebracht oder unangemessen. Die weiteren Anforderungen können ebenfalls erfüllt werden, nämlich die Bindung auf festgelegte Zwecke, und die Ausdrücklichkeit.

*Bewertung der Notwendigkeit und Angemessenheit*

Der Verarbeitungsvorgang 4 ist notwendig für seinen Zweck, da Versäumung des Termins eine große Gefahr ist, vor der die Geschäftsanbahnung gesichert werden muss. Mildere Mittel sind, zumal wegen der Freiwilligkeit, nicht erkennbar. Soweit, wie vorgesehen, rechtskonform eingewilligt wurde, sind auch die Interessen gleichgelagert, sodass die Angemessenheit zu bejahen ist.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Unter Verstoß gegen das Gewährleistungsziel der Integrität und der Verfügbarkeit könnte eine Terminerrinerung mit falschen Angaben oder nicht versendet werden und daher im Vertrauen auf die Erinnerung bzw. das Senden der Erinnerung ein Termin nicht oder nicht richtig wahrgenommen werden.

(b) Unter Verletzung des Ziels der Vertraulichkeit, Integrität und der Nichtverkettung könnte die Terminerinnerung an unberechtigte Personen gesendet werden, z. B. wenn eine falsche E-Mail-Adresse eingegeben oder ausgewählt wurde.

(c) Unter Verstoß gegen das Ziel der Datenminimierung könnte die Terminerinnerung unnötige Angaben enthalten, wodurch die Gefahr des Risikos b erhöht werden würde.

(d) Unter Verstoß gegen das Ziel der Intervenierbarkeit könnte die Freiwilligkeit nicht sichergestellt oder das tatsächliche Absenden der Erinnerung nicht kontrollierbar sein.

Bei eine Schwellwertanalyse ist nur das Kriterium der Betroffenheit von „vertraulichen oder höchstpersönlichen Daten“ vor. Zwar betreffen die Daten vorliegend die Intimsphäre, doch wird nur ein Bruchteil der Daten von 2. und 3. verarbeitet, weshalb das Risiko noch als **normal** eingestuft werden kann.

*Vorgeschlagene Abhilfemaßnahmen*

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Dem Risiko für die Integrität und Verfügbarkeit kann geeignet begegnet werden durch (1.) die Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen, (2.) die Festlegung des Sollverhaltens von Abläufen bzw. Prozessen und regelmäßiger Durchführung von Tests zur Feststellbarkeit bzw. Feststellung der Ist-Zustände von Prozessen, (3.) die Redundanz von Hard- und Software sowie Infrastruktur, (4.) die Umsetzung von Reparaturstrategien und Ausweichprozessen, sowie (5.) die Erstellung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit.

(b) Das Risiko für die Vertraulichkeit, Integrität und Nichtverkettung kann angemessen abgemildert werden durch (1.) die Festlegung und Kontrolle der Nutzung zugelassener Ressourcen insbesondere Kommunikationskanäle, (2.) das Löschen oder Berichtigen falscher Daten, (3.) Prozesse zur Aufrechterhaltung der Aktualität von Daten, (4.) eine Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Ne-

benwirkungen von Prozessen, (5.) eine Festlegung des Sollverhaltens von Abläufen bzw. Prozessen und regelmäßiger Durchführung von Tests zur Feststellbarkeit bzw. Feststellung der Ist-Zustände von Prozessen, sowie (6.) den Einsatz von zweckspezifischen Pseudonymen, Anonymisierungsdiensten, anonymen Credentials, Verarbeitung pseudonymer bzw. anonymisierter Daten.

(c) Dem Risiko für die Datenminimierung kann auf geeignete Weise begegnet werden (1.) durch die Reduzierung von Möglichkeiten der Kenntnisnahme vorhandener Daten, sowie (2.) durch die Reduzierung von erfassten Attributen der betroffenen Personen.

(d) Dem Risiko für die Intervenierbarkeit kann auf geeignete Weise begegnet werden (1.) durch Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten, sowie (2.) durch die Schaffung notwendiger Datenfelder z. B. für Sperrkennzeichen, Benachrichtigungen, Einwilligungen, Widersprüche, Gegendarstellungen.

Ergebnis:

Unter zweckmäßiger Kombination der bei 4.a bis 4.d aufgeführten Einzelmaßnahmen bieten sich für den Verarbeitungsvorgang 4 insgesamt die folgenden Maßnahmen an:

A) Festlegung des Sollverhaltens der Abläufe und Prozesse und Überprüfung dessen durch regelmäßige Durchführung von Software-Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen (4.a.1., 4.a.2., 4.b.4., 4.b.5.)

B) Herstellung von Redundanz von Hard- und Software sowie Infrastruktur sowie Festlegung und regelmäßige Kontrolle eines Verfahrens zur Wiederherstellung der Verarbeitungstätigkeit (4.a.3., 4.a.4., 4.a.5.)

C) Festlegung und Kontrolle der Nutzung zugelassener Ressourcen insbesondere Kommunikationskanäle (4.b.1.)

D) Einführung und Kontrolle von Prozessen zur Löschen oder Berichtigen falscher bzw. veralteter Daten (4.b.2., 4.b.3.)

E) Reduzierung der in der E-Mail enthaltenen Informationen zu dem Termin und der betroffenen Person auf ein das notwendige Minimum und, wo dies nicht möglich ist, die Verwendung von Pseudonymen oder vergleichbaren Informationen (4.b.6., 4.c.1., 4.c.2.)

F) Bereitstellung eines von den betroffenen Personen jederzeit, unkompliziert und ausschließlich bearbeitbaren Datenfeldes, das bestimmt, ob die Verarbeitung stattfindet (4.d.1., 4.d.2.)

**5. Vor-Ort-Identifizierung (Termine)**

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 5. | Vor-Ort-Identifizierung | Die Patient*innen erhalten über die Anwendung einen Code, mit dem sie sich ohne weitere Unterlagen in der Praxis vor Ort ausweisen können. Die Praxis kann den Code einem Termin und damit einer*m Patient*in zuordnen. | Wahrnehmung des Termins |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Auch Verarbeitungsvorgang 5 stützt sich, wie Nr. 2 und 3, aus den zuvor genannten Gründen auf die Art. 6 Abs. 1 Buchst b), Art. 9 Abs. 2 Buchst. h) DSGVO. Die Erforderlichkeit für den Vertrag ergibt sich hier daraus, dass die Wahrnehmung des Termins voraussetzt, dass die erschienene Person der Terminbuchung zugeordnet werden kann, und dass nur dadurch auch die adäquaten und gewünschten Behandlungsmaßnahmen vorgenommen werden können. Damit ist die Verarbeitung für die Erfüllung des Behandlungsvertrags unerlässlich.

*Bewertung der Notwendigkeit und Angemessenheit*

Der Verarbeitungsvorgang 5 ist notwendig für seinen Zweck, denn irgendeine Form der Identifizierung ist – sogar objektiv – notwendig und mildere Formen als die hier angeführte sind nicht ersichtlich. Da beide Beteiligte ein Interesse an der richtigen Wahrnehmung des Termins haben, ist die Verarbeitung zudem angemessen.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Es besteht das Risiko, dass unter Nichtbeachtung der Gewährleistungsziele der Verfügbarkeit und der Integrität der bereitgestellte Code nicht funktioniert, nicht akzeptiert wird, fehlerhaft ist oder als ungültig angesehen wird und dadurch der Termin nicht wahrgenommen werden kann.

(b) Unter Verstoß gegen die Gewährleistungsziele der Transparenz, der Vertraulichkeit und der Nichtverkettung könnte nicht eindeutig bestimmt sein, wer einen Code zuordnen und damit Zugriff auf die Terminsinformationen erhalten kann, wodurch unbefugte Personen Zugriff auf diese erhalten könnten.

(c) Dem Ziel der Verfügbarkeit und Intervenierbarkeit zuwider könnte es keine Möglichkeit geben, den Code zu widerrufen oder einen neuen Code auszustellen, wenn der alte Code vergessen oder verloren wurde.

Bei einer Schwellwertanalyse erweist sich die Kriterien der Betroffenheit von „vertraulichen oder höchstpersönlichen Daten“ als anwendbar. Zudem können durch Fehler der Verarbeitung „Betroffene an der Ausübung eines Rechts oder der Nutzung einer Dienstleistung bzw. Durchführung eines Vertrags gehindert“ werden, wobei dieses Kriterium nur als geringfügig betroffen anzusehen ist, da es den Praxen freisteht, die Person auch auf andere Weise zu identifizieren und somit ein Fehler der Verarbeitung nicht zwangsläufig und wenn dann nur mittelbar zu der Folge führt. Daher ist das Risiko noch als **normal** einzustufen.

*Vorgeschlagene Abhilfemaßnahmen*

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Dem Risiko für die Verfügbarkeit und Integrität kann angemessen begegnet werden (1.) durch die Umsetzung von Reparaturstrategien und Ausweichprozessen, (2.) durch die Erstellung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit, (3.) die Garantierung alternativer Identifikationsmittel, (4.) eine Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen, sowie (5.), eine Festlegung des Sollverhaltens von Abläufen bzw. Prozessen und regelmäßiger Durchführung von Tests zur Feststellbarkeit bzw. Feststellung der Ist-Zustände von Prozessen.

(b) Das Risiko für die Transparenz, die Vertraulichkeit und die Nichtverkettung kann vorliegend geeignet (1.) durch Dokumentation der Verträge mit den internen Mitarbeitenden, Verträge mit externen Dienstleistern und Dritten, von denen Daten erhoben bzw. an die Daten übermittelt werden, Geschäftsverteilungspläne, Zuständigkeitsregelungen, (2.) durch die Protokollierung von Zugriffen und Änderungen, (3.) durch Dokumentation der Verarbeitungsprozesse mittels Protokollen auf der Basis eines Protokollierungs- und Auswertungskonzepts, (4.) durch die Benachrichtigung von Betroffenen bei Datenpannen oder bei Weiterverarbeitungen zu einem anderen Zweck, (5.) durch die Festlegung eines Berechtigungs- und Rollenkonzeptes nach dem Erforderlichkeitsprinzip auf der Basis eines Identitätsmanagements durch die verantwortliche Stelle, (6.) durch die Eingrenzung der zulässigen Personalkräfte auf solche, die nachprüfbar zuständig (örtlich, fachlich), fachlich befähigt, zuverlässig (ggf. sicherheitsüberprüft) und formal zugelassen sind sowie keine Interessenskonflikte bei der Ausübung aufweisen, (7.) durch Festlegung und Kontrolle der Nutzung zugelassener Ressourcen insbesondere Kommunikationskanäle, (8.), durch Festlegung und Kontrolle organisatorischer Abläufe, interner Regelungen und vertraglicher Verpflichtungen (Verpflichtung auf Datengeheimnis, Verschwiegenheitsvereinbarungen usw.), (9.) durch Verschlüsselung von gespeicherten oder transferierten Daten sowie Prozesse zur Verwaltung und zum Schutz der kryptografischen Informationen, (10.) durch die

Trennung nach Organisations-/Abteilungsgrenzen, sowie (11.) durch den Einsatz von zweckspezifischen Pseudonymen, Anonymisierungsdiensten, anonymen Credentials, Verarbeitung pseudonymer bzw. anonymisierter Daten.

(c) Dem Risiko für die Verfügbarkeit und Intervenierbarkeit kann angemessen (1.) durch die Umsetzung von Reparaturstrategien und Ausweichprozessen, sowie (2.) durch eine Möglichkeit der Wiederherstellung personenbezogener Daten begegnet werden.

Ergebnis:

Unter zweckmäßiger Kombination der bei 5.a bis 5.c aufgeführten Einzelmaßnahmen bieten sich für den Verarbeitungsvorgang 5 insgesamt die folgenden Maßnahmen an:

A) Gewährleistung des Praxisbetriebs im Falle einer (technischen) Störung der Plattform oder der lokalen Informationstechnik durch physische Kopien (5.a.1., 5.a.2.)

B) Gewährleistung der Wahrnahme von Terminen im Falle eines vergessenen oder nichtfunktionierenden Buchungscodes durch Bereithaltung alternativer Identifizierungsverfahren sowie einer Möglichkeit zur Wiederherstellung des Buchungscodes (5.a.1., 5.a.3., 5.c.1., 5.c.2.)

C) Konzeption und Erprobung einer tatsächlich umsetzbaren Reparaturstrategie mittels Redundanz der Daten, Systeme und Prozesse zur Vermeidung und Reduktion der Auswirkungen von Systemausfällen (5.a.1., 5.a.2., 5.c.1.)

D) Festlegung des Sollverhaltens der Abläufe und Prozesse und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen (5.a.4., 5.a.5., 5.b.3.)

E) Dokumentation bzw. Protokollierung der Erteilung, Änderung und Rücknahme von Zugriffsberechtigungen (inkl. damit verbundener Verträge bzw. Vereinbarungen/ Weisungen/Geschäftsverteilungspläne bzw. Zuständigkeitsregelungen/Belehrungen) und tatsächlichen Zugriffen/Verarbeitungsvorgängen sowie automatisierte Benachrichtigung bei unberechtigten/verdächtigen Vorgängen (5.b.1., 5.b.2., 5.b.3., 5.b.8.)

F) Benachrichtigung von Betroffenen bei festgestellten unberechtigten Zugriffen (5.b.4.)

G) Beschränkung des Zugriffs auf gesicherte und gehärtete, explizit individuell zugelassene Computer, die sich physisch in der Praxis befinden, sowie auf ausdrücklich individuell authentifizierte, explizit autorisierte, förmlich verpflichtete, zu einer bestimmten Verhaltensweise geschulte und dienstlich angewiesene und vertrauenswürdige Personalkräfte (5.b.5., 5.b.6., 5.b.7., 5.b.8.)

H) Einsatz von Transportverschlüsselung von Daten zwischen der Plattform und dem Praxissystem (5.b.9.)

I) Trennung nach Organisations-/Abteilungsgrenzen, insbesondere bei Gemeinschaftspraxen oder bei Zweig- bzw. weiteren Praxen eines Arztes oder einer Ärztin oder bei verschiedenen trennbaren Tätigkeiten (5.b.10.)

J) Verwendung von Pseudonymen dort, wo es nicht auf die Identität der Person ankommt oder die Identität auf einem anderen Wege sichergestellt werden kann (5.b.11.)

## 6. Erzeugung von Trainingsdaten (Chatbot)

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 6. | Erzeugung von Trainingsdaten | Für das Training eines auf „KI“ basierenden Chatbots, der versucht, im Dialog medizinische Probleme zu diagnostizieren und entweder eine bestimmte Ärztin/einen bestimmten Arzt vorschlägt oder beruhigt, dass es nur eine milde Erkrankung sei, müssen exemplarische Chatverläufe erzeugt werden. Hierzu wird ein Live-Chat zwischen (echten) Patient*innen und (echten) Ärzt*innen erzeugt und aufgezeichnet. Am Ende wählen die Ärzt*innen die richtige Antwort aus. | Ermöglichung/Verbesserung des Chatbots |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Als Rechtsgrundlage für die Erzeugung (und Verwendung) von Trainingsdaten kommt die Einwilligung (Art. 6 Abs. 1 Buchst. a) DSGVO) nicht in Betracht. Zwar ist die Möglichkeit einer freiwillige Abgabe der Einwilligung nicht von vorne herein auszuschließen, jedoch kann die Widerruflichkeit der Einwilligung nicht gewährleistet werden. Sobald personenbezogene Daten in einem „KI“-System eingegeben sind, können diese mit ausreichender Wahrscheinlichkeit wieder hervorgerufen werden. Da bei einem Widerruf die Verarbeitung ex nunc aufzuhören hätte, müsste es ausgeschlossen werden, dass ab dem Zeitpunkt des Widerrufs personenbezogene Daten hervorgerufen werden können. Dies könnte jedoch mit der datenschutzrechtlich erforderlichen Sicherheit nur dadurch erfolgen, dass das System neu – diesmal ohne die jeweiligen Daten – trainiert würde, was mit einem unangemessen hohen Aufwand verbunden wäre und daher für die Verantwortlichen nicht realistisch in Betracht kommt.

Auch die Rechtsgrundlage der Erforderlichkeit für die Erfüllung eines Vertrags kommt nicht in Betracht. Für einen Behandlungs-/Diagnosevertrag, der mit dem aufgezeichneten Chat erfüllt wird, ist die Verarbeitung schon objektiv nicht erforderlich. Ein „Datenkaufvertrag“ kommt ebenfalls nicht in Betracht, da Art. 6 Abs. 1 Buchst. b) DSGVO die Verarbeitung personenbezogener Daten zur Erfüllung eines Vertrags, nicht jedoch als Erfüllung eines Vertrags vorsieht. Die gesetzlichen Anforderungen an eine Einwilligung sollen nicht durch eine einwilligungsgleiche Gestaltung als Vertrag umgangen werden können.

Als Rechtsgrundlage kommen jedoch die berechtigten Interessen des Verantwortlichen in Betracht. Das Interesse daran, einen Chatbot zu trainieren, wird vorliegend auch nicht von dem Interesse an der Nichtverwendung überwogen, wenn die Teilnahme an dem Chat freiwillig erfolgt und dem Risiko, gerade wegen der weitestgehenden Unmöglichkeit der Entfernung der Daten aus dem Modell, angemessen Rechnung getragen wird (s.u.).

Vorliegend sind zudem offensichtlich Gesundheitsdaten betroffen, sodass ein Ausnahmegrund nach Art. 9 Abs. 2 DSGVO vorliegen muss. Aus den gleichen Gründen wie zuvor kommen weder die Einwilligung (Art. 9 Abs. 2 Buchst. a) DSGVO) noch die Erforderlichkeit für einen Diagnose- oder Behandlungsvertrag (Art. 9 Abs. 2 Buchst. h) DSGVO) in Betracht.

Auch der Ausnahmegrund des Art. 9 Abs. 2 Buchst. j) DSGVO in Verbindung mit § 27 BDSG kann das Verarbeitungsverbot nicht aufheben. Ungeachtet dessen, ob es sich bei dem Trainieren eines „KI"-Modells überhaupt um wissenschaftliche Forschung handelt oder ob das Interesse des Verantwortlichen hier die Interessen der betroffenen Person „erheblich überwiegen", liegt das Tatbestandsmerkmal der Verarbeitung „im öffentlichen Interesse" des Art. 9 Abs. 2 Buchst. j) DSGVO nicht vor. Die Verarbeitung dient nämlich alleine den privaten wirtschaftlichen Interessen des Verantwortlichen.

Art. 9 Abs. 2 Buchst. g) DSGVO in Verbindung mit § 27 BDSG ist ebenfalls nicht anwendbar, denn die Verarbeitung ist nicht für § 27 BDSG „erforderlich". Ausnahmegrund g) betrifft nämlich ausschließlich gesetzliche Verpflichtungen, während § 27 BDSG nur eine gesetzliche Berechtigung enthält.

Aus den bereits genannten Gründen sind schließlich auch alle anderen Ausnahmegründe des Art. 9 Abs. 2 DSGVO nicht einschlägig. Die Verarbeitung verstößt mithin gegen das Verbot des Art. 9 Abs. 1 DSGVO und ist **rechtswidrig**.

*Bewertung der Notwendigkeit und Angemessenheit*

n. a.

*Bewertung der Risiken für die betroffenen Personen*

n. a.

*Vorgeschlagene Abhilfemaßnahmen*

Die Verarbeitung ist rechtswidrig und darf daher nicht vorgenommen werden.

### 7. Training (Chatbot)

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 7. | Training | Die aufgezeichneten Chatverläufe werden in der „KI“ trainiert, bis diese plausibel klingende Chatverläufe erzeugt. | Ermöglichung/Verbesserung des Chatbots |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Für Verarbeitungsvorgang 7 kann aus den dort angeführten Gründen nichts anderes gelten als für Verarbeitungsvorgang 6. Zwar liegt mit Art. 6 Abs. 1 Buchst. f) eine tragfähige Rechtsgrundlage für die Verarbeitung vor, doch vermag das Verarbeitungsverbot des Art. 9 Abs. 1 nicht aufgehoben werden, sodass die Verarbeitung ebenfalls **rechtswidrig** ist.

*Bewertung der Notwendigkeit und Angemessenheit*

n. a.

*Bewertung der Risiken für die betroffenen Personen*

n. a.

*Vorgeschlagene Abhilfemaßnahmen*

Die Verarbeitung ist rechtswidrig und darf daher nicht vorgenommen werden.

## 8. Konsultation (Chatbot)

*Systematische Beschreibung der Verarbeitungstätigkeit*

| Nr. | Vorgang | Beschreibung | Zwecke |
|---|---|---|---|
| 8. | Konsultation | Jeder Person, die einen Termin buchen möchte, wird durch die Plattform dringend empfohlen, zunächst den Chatbot zu nutzen, bevor ein Termin gebucht wird. Sie beantworten ein paar Fragen über ihr Anliegen bzw. medizinisches Problem und erhalten dann entweder direkt eine Ärztin oder einen Arzt vorgeschlagen oder eine beruhigende Diagnose. | Vereinfachung für Patient*innen<br><br>Entlastung des med. Personals |

*Rechtmäßigkeit nach den Art. 6 und 9 DSGVO*

Verarbeitungsvorgang 8 kann – jedoch nur mit Auflagen – auf die Rechtsgrundlage der Einwilligung (Art. 6 Abs. 1 Buchst. a) DSGVO) gestützt werden. Die Einwilligung wäre jedoch nur wirksam, wenn vor Verwendung des Chatbots über die Zwecke der Verarbeitung und ihre Risiken, insbesondere die Unrichtigkeit des Konsultationsergebnisses informiert wird. Zudem kann die Freiwilligkeit nur vorliegen, wenn von der „dringenden Empfehlung“ Abstand genommen wird. Jedwedes Drängen oder Irreführen dürfte zur Unwirksamkeit der Einwilligung und damit zur Rechtswidrigkeit des Verarbeitungsvorgangs führen. Zudem ist sicherzustellen, dass die Einwilligung jederzeit widerrufen werden kann, sodass die Verwendung von Chatverläufen zum Training aus den bei Verarbeitungsvorgang 6 beschriebenen Gründen nicht stattfinden darf.

Soweit diese Anforderungen erfüllt sind, kann die Einwilligung auch als Ausnahmegrund im Sinne des Art. 9 Abs. 2 Buchst. a) DSGVO in Betracht kommen. Zu beachten ist, dass die Einwilligung „ausdrücklich“ zu erfolgen hat. Die Verwendung des Chatbots, d.h. schlüssiges Verhalten, genügt nicht, um eine Einwilligung anzunehmen.

*Bewertung der Notwendigkeit und Angemessenheit*

Der Verarbeitungsvorgang ist noch gerade so als geeignet für seine Zwecke anzusehen. Soweit die „KI“ nicht hinreichend zuverlässig ist, führt sie zwar weder zu einer Vereinfachung für die Patient*innen noch (sowohl kurz- wie mittelfristig) für eine Entlastung des medizinischen Personals, jedoch kann im Falle expliziter Belehrung und dennoch erfolgter Einwilligung durch die betroffenen Personen das Begehren nach einer wenigstens ungefähren Einschätzung angenommen werden, womit die Eignung vorliegen würde.

Es gibt zwar mit einem telefonischen Bereitschaftsdienst ein zumutbares milderes Mittel, jedoch kann dies wegen des hohen Personalaufwands als weniger geeignet angesehen werden. Andere Formen der Kommunikation, die synchron sind, leiden unter dem gleichen Problem und sind, wenn sie asynchron sind, wegen der damit verbundenen Wartezeiten ungeeignet. Damit sind wenigstens gleich geeignete, mildere Mittel nicht ersichtlich.

Die Verarbeitung kann nur mit den bereits bei der Rechtmäßigkeit verlangten Auflagen als angemessen angesehen werden, d.h. nur dann, wenn sie freiwillig und nach einer umfassenden Aufklärung von der betroffenen Person gewünscht wird. In allen anderen Fällen würden die Interessen der betroffenen Personen an wirkungsvoller medizinischer Begutachtung die potentiellen Interessen an dem Betrieb des Chatbots weit überragen.

*Bewertung der Risiken für die betroffenen Personen*

Es können die folgenden Risiken identifiziert werden:

(a) Unter Verstoß gegen das Gewährleistungsziel der Integrität könnte der Chatbot falsche Antworten liefern, was zu schwerwiegenden medizinischen Folgen für die betroffende Person wegen falscher Diagnostik, bis hin zum Tod, führen könnte.

(b) Unter Verletzung der Nichtverkettung und Vertraulichkeit könnten dem Chatbot eröffnete Informationen über das Gesundheitsproblem/Anliegen gespeichert oder in den Chatbot eingepflegt werden, sodass sie von Unbefugten zur Kenntnis genommen oder im Gespräch mit dem Chatbot repliziert werden könnten.

(c) Unter Verstoß gegen den Grundsatz der Transparenz könnte der Chatbot nicht hinreichend nachvollziehbar sein, sodass nicht ausgeschlossen werden kann, dass sich das Risiko a oder b verwirklicht.

(d) Unter Nichtgewährleistung der Verfügbarkeit könnte der Chatbot nicht verfügbar sein, sodass dringende medizinische Fragen nicht beantwortet werden, da sich die betroffene Person auf die Antworten des Chatbots verlässt.

(e) Unter Missachtung des Grundsatzes der Datenminimierung könnte der Chatbot Fragen nach nicht unbedingt erforderlichen Gesundheitsdaten stellen, sodass in unangemessener Weise in die Rechte und Freiheiten der betroffenen Person eingegriffen werden und das Risiko b erhöht würde.

(f) Unter Verstoß gegen die Intervenierbarkeit und Datenminimierung könnte es nicht möglich sein, in den Chatbot eingegebene Informationen zu löschen oder die Einwilligung zu widerrufen, insbesondere wenn die Informationen in den Chatbot eingepflegt wurden.

(g) Unter Verstoß gegen die Intervenierbarkeit könnte es nicht möglich sein, ohne Inanspruchnahme des Chatbots einen Termin zu buchen.

Bei einer Schwellwertanalyse erweisen sich die Kriterien der Betroffenheit „vertraulicher oder höchstpersönlicher Daten“, der „Datenverarbeitung in großem Umfang“, des „Abgleichens oder Zusammenführens von Datensätzen“ sowie der „Innovativen Nutzung oder Anwendung neuer technologischer oder organisatorischer Lösungen“ als zutreffend. Alleine schon aus diesem Grund ist das Risiko **hoch** einzustufen.

*Vorgeschlagene Abhilfemaßnahmen*

Damit eine Einwilligung in die Verwendung der eigenen Daten im Rahmen dieses Verarbeitungsvorgangs in Betracht kommt, muss der beabsichtigte Verarbeitungsvorgang so modifiziert werden, dass die Verwendung des Chatbots echt freiwillig ist. Der Chatbot darf neutral als Möglichkeit präsentiert werden, jedoch darf weder ein Drängen noch ein Nudgen verwendet werden, noch darf der Eindruck erweckt werden, dass die Verwendung vorgeschrieben ist, noch darf die Verwendung tatsächlich vorgeschrieben werden. Zudem ist auf die Unzuverlässigkeit und die Risiken der Verwendung explizit, ausdrücklich und in verständlicher Weise hingewiesen werden, bevor die Einwilligung abgefragt wird.

Aus den zuvor aufgeführten und untersuchten Risiken ergeben sich folgende Abhilfemaßnahmen:

(a) Dem Risiko für die Integrität kann vorliegend geeignet teilweise begegnet werden durch (1.) das Löschen oder Berichtigen falscher Daten – insbesondere bei den Trainingsdaten –, (2.) Prozesse zur Aufrechterhaltung der Aktualität von Daten, sowie (3.) eine Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen.

(b) Das Risiko für die Nichtverkettung und die Vertraulichkeit kann geeignet abgemildert werden durch (1.) die programmtechnische Unterlassung bzw. Schließung von Schnittstellen bei Verarbeitungsverfahren und Komponenten, (2.) regelnde Maßgaben zum Verbot von Backdoors sowie qualitätssichernde Revisionen zur Compliance bei der Softwareentwicklung, (3.) die Verschlüsselung von gespeicherten oder transferierten Daten sowie Prozesse zur Verwaltung und zum Schutz der kryptografischen Informationen, (4.) den Schutz vor äußeren Einflüssen, sowie (5.) die Festlegung und Kontrolle organisatorischer Abläufe, interner Regelungen und vertraglicher Verpflichtungen (Verpflichtung auf Datengeheimnis, Verschwiegenheitsvereinbarungen usw.).

(c) Dem Risiko für die Transparenz kann angemessen teilweise begegnet werden durch (1.) Dokumentation der Bestandteile von Verarbeitungstätigkeiten insbesondere der Geschäftsprozesse, Datenbestände, Datenflüsse und Netzpläne, dafür genutzte IT-Systeme, Betriebsabläufe, Beschreibungen von Verarbeitungstätigkeiten, Zusammenspiel mit anderen Verarbeitungstätigkeiten, (2.) Dokumentation von Tests, der Freigabe und ggf. der Datenschutz-Folgenabschätzung von neuen oder geänderten Verarbeitungstätigkeiten, sowie (3.) Dokumentation der Faktoren, die für eine Profilierung, zum Scoring oder für teilautomatisierte Entscheidungen genutzt werden.

(d) Dem Risiko für die Verfügbarkeit kann angemessen begegnet werden durch (1.) den Schutz vor äußeren Einflüssen, (2.) die Redundanz von Hard- und Software sowie Infrastruktur, (3.) die Umsetzung von Reparaturstrategien und Ausweichprozessen, sowie (4.) die Erstellung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit.

(e) Das Risiko für die Datenminimierung kann (1.) durch Festlegung von Voreinstellungen – bzw. Trainingsparametern – für betroffene Personen, die die Verarbeitung ihrer Daten auf das für den Verarbeitungszweck erforderliche Maß beschränken, (2.) durch Festlegung und Umsetzung eines Löschkonzepts, sowie (3.) durch Regelungen zur Kontrolle von Prozessen zur Änderung von Verarbeitungstätigkeiten teilweise abgemildert werden.

(f) Dem Risiko für die Intervenierbarkeit und Datenminimierung kann geeignet abgemildert werden durch (1.) Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten, (2.) die Schaffung notwendiger Datenfelder z. B. für Sperrkennzeichen, Benachrichtigungen, Einwilligungen, Widersprüche, Gegendarstellungen, (3.) die Implementierung von Deaktivierungsmöglichkeit einzelner Funktionalitäten ohne Mitleidenschaft für das Gesamtsystem, sowie (4.) Einführung von Regelungen zur Kontrolle von Prozessen zur Änderung von Verarbeitungstätigkeiten.

(g) Dem Risiko für die Intervenierbarkeit kann angemessen begegnet werden durch (1.) Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten, (2.) die Implementierung von Deaktivierungsmöglichkeit einzelner Funktionalitäten ohne Mitleidenschaft für das Gesamtsystem, sowie (3.) das Bereitstellen von Optionen für Betroffene, um Programme datenschutzgerecht einstellen zu können.

Ergebnis:

Unter zweckmäßiger Kombination der bei 8.a bis 8.g aufgeführten Einzelmaßnahmen sowie der Auflagen aus der Rechtmäßigkeitskontrolle bieten sich für den Verarbeitungsvorgang 5 insgesamt die folgenden Maßnahmen an:

A) Anzeige einer verständlichen Warnung vor der Unzuverlässigkeit und den Risiken der Verwendung des Chatbots, bevor eine Einwilligung eingeholt wird

B) Einführung und Umsetzung von Prozessen zur regelmäßigen Löschung oder Berichtigung falscher oder veralteter Trainingsdaten (8.a.1., 8.a.2., 8.b.4.)

C) Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen (8.a.3.)

D) Dokumentation der Bestandteile von Verarbeitungstätigkeiten insbesondere der Geschäftsprozesse, Datenbestände, Datenflüsse und Netzpläne, dafür genutzte IT-Systeme, Betriebsabläufe, Beschreibungen von Verarbeitungstätigkeiten, Zusammenspiel mit anderen Verarbeitungstätigkeiten (8.c.1.)

E) Dokumentation der Maßnahmen zur Gewährleistung einer hohen Datenqualität und einer hohen Zuverlässigkeit, insbesondere Dokumentation von Tests und der Freigabe sowie Dokumentation der Faktoren, die von dem Chatbot – mutmaßlich – verwendet werden (8.c.2., 8.c.3.)

F) Programmtechnische Unterlassung bzw. Schließung von Schnittstellen bei Verarbeitungsverfahren und Komponenten und Gewährleistung des Erhalts dieser Maßnahme durch Regelnde Maßgaben zum Verbot von Backdoors sowie qualitätssichernde Revisionen zur Compliance bei der Softwareentwicklung (8.b.1., 8.b.2.)

G) Verschlüsselung von gespeicherten oder transferierten Daten sowie Prozesse zur Verwaltung und zum Schutz der kryptografischen Informationen (8.b.3.)

H) Festlegung und Umsetzung eines Löschkonzepts (8.e.2.)

I) Schutz vor äußeren Einflüssen, insbesondere vor unberechtigtem Zugriff auf oder unberechtigter Veränderung der Daten oder Prozesse (8.b.4., 8.d.1.)

J) Festlegung und Kontrolle organisatorischer Abläufe, interner Regelungen und vertraglicher Verpflichtungen (Verpflichtung auf Datengeheimnis, Verschwiegenheitsvereinbarungen usw.) (8.b.5.)

K) Erstellung und regelmäßige Überprüfung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit auf Basis der Redundanz von Hard- und Software sowie Infrastruktur (8.d.4., 8.d.2.)

L) Festlegung von Voreinstellungen und Daten/Parametern beim Training, die die Verarbeitung der personenbezogenen Daten auf das für den Verarbeitungszweck erforderliche Maß beschränken (8.e.1.)

M) Regelungen zur Kontrolle von Prozessen zur Änderung von Verarbeitungstätigkeiten (8.e.3., 8.f.4.)

N) Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten und Schaffung notwendiger Datenfelder zur Umsetzung dieser Maßnahmen (8.f.1., 8.g.1., 8.f.2.)

O) Gewährleistung der Möglichkeit, die Hauptanwendung zu nutzen, wenn der Chatbot nicht funktioniert, die Einwilligung verweigert oder widerrufen oder die Nutzung auf andere Weise abgelehnt wird (8.f.3., 8.g.2., 8.d.3., 8.g.3.)

Da für die Verarbeitungstätigkeit ein hohes Risiko festgestellt worden ist, müssen diese Maßnahmen besonders wirksam umgesetzt werden, z. B. muss bei 8.G. eine dem Stand der Technik entsprechende und nach den gegenwärtigen Erkenntnissen auch auf absehbare Zeit sichere Verschlüsselungstechnik verwendet werden oder bei 8.C. auf eine hohe Testdichte geachtet werden.

Zudem ist zu berücksichtigen, dass den hohen Risiken wegen der inhärenten Unzuverlässigkeit eines „KI“-Systems nur teilweise geeignet begegnet werden konnte. Daher ist die **Vorherige Konsultation der Aufsichtsbehörde** nach Art. 36 DSGVO zwingend erforderlich, bevor mit der Durchführung von Verarbeitungsvorgang 8 begonnen wird.

**Ergebnisse der Datenschutz-Folgenabschätzung**

Aufgrund der Ergebnisse der Datenschutz-Folgenabschätzung müssen die Verarbeitungsvorgänge wie folgt angepasst werden, um mit dem Datenschutz im Einklang zu stehen:

1. Für die Verarbeitungsvorgänge 2 bis 5 wurde die Verantwortlichkeit geklärt, sodass die Verarbeitung nunmehr unter der Verantwortung des medizinischen Fachpersonals erfolgt und das Unternehmen ausschließlich als Auftragsverarbeiter tätig wird.

2. Die Verarbeitungsvorgänge 6 und 7 wurden gestrichen, da sie nicht rechtmäßig umgesetzt werden können.

3. Die in Anlage 1 noch einmal aufgelisteten 52 technisch-organisatorischen Maßnahmen werden umgesetzt.

4. Bezüglich den Verarbeitungsvorgang 8 ist die Aufsichtsbehörde vor Beginn der Verarbeitung zu konsultieren, da das hohe Risiko nicht ausreichend abgemildert werden konnte.

# Anlage 1
# Technisch-Organisatorische Maßnahmen

| # | Adresse | Beschreibung der Maßnahme |
|---|---|---|
| **1. Anzeige von Arztinformationen** | | |
| 1 | **1.A.** | Möglichkeit zum Ändern oder Löschen von eingetragenen Daten |
| 2 | **1.B.** | Festlegung und Umsetzung eines Konzepts mit Prozessen zur Löschung oder Berichtigung falscher oder veralteter Daten |
| 3 | **1.C.** | Dokumentation der Faktoren und Bestandteile der Darstellung von Ärzte-profilen |
| 4 | **1.D.** | Beschränkung der (standardmäßig) erfassten Attributen bzgl. Ärzt*innen und Arztpraxen auf den notwendigen Inhalt |
| **2. Buchung eines Termins – HOHES RISIKO!** | | |
| 5 | **2.A.** | Beschränkung des Zugriffs zu den Terminen und ihre Verarbeitung auf au-thentifizierte und (anhand ihrer Prozessrolle sowie im Einzelfall) berechtig-te Personen |
| 6 | **2.B.** | Einführung von Anforderungen an Personen (außer den Patient*innen), die auf Termineinträge zugreifen wollen, wie bspw. nachprüfbare Zuständigkeit, fachliche Befähigung, Zuverlässigkeit oder die förmliche Verpflichtung zur Verschwiegenheit |
| 7 | **2.C.** | Beschränkung des (technischen wie physischen) Zugangs und Schutz des Rechenzentrums, in dem die Verarbeitungstätigkeit stattfindet, wie z. B. durch eine Firewall, ein Schadsoftware-Monitoring oder eine Sicherheits-schleuse |
| 8 | **2.D.** | Transportverschlüsselung bei Anfragen an die Plattform, Verschlüsselung der gespeicherten Daten, sowie Einführung eines Kryptokonzepts |
| 9 | **2.E.** | Logische und ggf. physische Trennung der Daten nach Organisations-/Ab-teilungsgrenzen |
| 10 | **2.F.** | Redundanz von Hard- und Software sowie Infrastruktur sowie Festlegung und regelmäßige Kontrolle eines Verfahrens zur Wiederherstellung der |

| | | |
|---|---|---|
| | | Verarbeitungstätigkeit |
| 11 | **2.G.** | Reduktion der standardmäßig erfassten Angaben bei der Terminbuchung auf notwendige Angaben wie Zeitpunkt und ausgewählte Praxis |
| 12 | **2.H.** | Protokollierung von Zugriffen und Änderungen sowie Benachrichtigung bei unberechtigten Maßnahmen |
| 13 | **2.I.** | Dokumentation der Möglichkeiten, wie auf die Termindaten Zugriff erhalten werden kann, welche Schutzmaßnahmen dagegen bestehen und wie ein (berechtigter oder unberechtigter) Zugriff nachvollzogen werden kann |
| 14 | **2.J.** | Möglichkeit der Änderung und Löschung von eingetragenen Daten durch die betroffenen Personen |
| **3. Verwaltung eines Termins – HOHES RISIKO!** | | |
| 15 - 19 | **3.A.** | *(die Maßnahmen 2.A, 2.B, 2.C, 2.D und 2.E)* |
| 20 | **3.B.** | Dokumentation der Zuweisung von Berechtigungen und Rollen |
| 21 | **3.C.** | Entwicklung, Dokumentation und Umsetzung eines Konzepts zur Protokollierung von Zugriffen und Änderung inkl. einer (wiederherstellbaren) Versionierung der Daten |
| **4. Terminerinnerung** | | |
| 22 | **4.A.** | Festlegung des Sollverhaltens der Abläufe und Prozesse und Überprüfung dessen durch regelmäßige Durchführung von Software-Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen |
| 23 | **4.B.** | Redundanz von Hard- und Software sowie Infrastruktur sowie Festlegung und regelmäßige Kontrolle eines Verfahrens zur Wiederherstellung der Verarbeitungstätigkeit |
| 24 | **4.C.** | Festlegung und Kontrolle der Nutzung zugelassener Ressourcen insbesondere Kommunikationskanäle |
| 25 | **4.D.** | Einführung und Kontrolle von Prozessen zur Löschen oder Berichtigen fal- |

| | | |
|---|---|---|
| | | scher bzw. veralteter Daten |
| 26 | **4.E.** | Reduzierung der in der E-Mail enthaltenen Informationen zu dem Termin und der betroffenen Person auf ein das notwendige Minimum und, wo dies nicht möglich ist, die Verwendung von Pseudonymen oder vergleichbaren Informationen |
| 27 | **4.F.** | Bereitstellung eines von den betroffenen Personen jederzeit, unkompliziert und ausschließlich bearbeitbaren Datenfeldes, das bestimmt, ob die Verarbeitung stattfindet |
| **5. Vor-Ort-Identifizierung** | | |
| 28 | **5.A.** | Gewährleistung des Praxisbetriebs im Falle einer (technischen) Störung der Plattform oder der lokalen Informationstechnik durch physische Kopien |
| 29 | **5.B.** | Gewährleistung der Wahrnahme von Terminen im Falle eines vergessenen oder nichtfunktionierenden Buchungscodes durch Bereithaltung alternativer Identifizierungsverfahren sowie einer Möglichkeit zur Wiederherstellung des Buchungscodes |
| 30 | **5.C.** | Konzeption und Erprobung einer tatsächlich umsetzbaren Reparaturstrategie mittels Redundanz der Daten, Systeme und Prozesse zur Vermeidung und Reduktion der Auswirkungen von Systemausfällen |
| 31 | **5.D.** | Festlegung des Sollverhaltens der Abläufe und Prozesse und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen |
| 32 | **5.E.** | Dokumentation bzw. Protokollierung der Erteilung, Änderung und Rücknahme von Zugriffsberechtigungen (inkl. damit verbundener Verträge bzw. Vereinbarungen/Weisungen/Geschäftsverteilungspläne bzw. Zuständigkeitsregelungen/Belehrungen) und tatsächlichen Zugriffen/Verarbeitungsvorgängen sowie automatisierte Benachrichtigung bei unberechtigten/verdächtigen Vorgängen |
| 33 | **5.F.** | Benachrichtigung von Betroffenen bei festgestellten unberechtigten Zugriffen |
| 34 | **5.G.** | Beschränkung des Zugriffs auf gesicherte und gehärtete, explizit individuell zugelassene Computer, die sich physisch in der Praxis befinden, sowie auf |

| | | |
|---|---|---|
| | | ausdrücklich individuell authentifizierte, explizit autorisierte, förmlich verpflichtete, zu einer bestimmten Verhaltensweise geschulte und dienstlich angewiesene und vertrauenswürdige Personalkräfte |
| 35 | **5.H.** | Transportverschlüsselung von Daten zwischen der Plattform und dem Praxissystem |
| 36 | **5.I.** | Trennung nach Organisations-/Abteilungsgrenzen, insbesondere bei Gemeinschaftspraxen oder bei Zweig- bzw. weiteren Praxen eines Arztes oder einer Ärztin oder bei verschiedenen trennbaren Tätigkeiten |
| 37 | **5.J.** | Verwendung von Pseudonymen dort, wo es nicht auf die Identität der Person ankommt oder die Identität auf einem anderen Wege sichergestellt werden kann |
| **6. Erzeugung von Trainingsdaten – rechtswidrig** | | |
| **7. Training – rechtswidrig** | | |
| **8. Konsultation – HOHES RISIKO!** | | |
| 38 | **8.A.** | Anzeige einer verständlichen Warnung vor der Unzuverlässigkeit und den Risiken der Verwendung des Chatbots, bevor eine Einwilligung eingeholt wird |
| 39 | **8.B.** | Einführung und Umsetzung von Prozessen zur regelmäßigen Löschung oder Berichtigung falscher oder veralteter Trainingsdaten |
| 40 | **8.C.** | Festlegung des Sollverhaltens von Prozessen und regelmäßiger Durchführung von Tests zur Feststellung und Dokumentation der Funktionalität, von Risiken sowie Sicherheitslücken und Nebenwirkungen von Prozessen |
| 41 | **8.D.** | Dokumentation der Bestandteile von Verarbeitungstätigkeiten insbesondere der Geschäftsprozesse, Datenbestände, Datenflüsse und Netzpläne, dafür genutzte IT-Systeme, Betriebsabläufe, Beschreibungen von Verarbeitungstätigkeiten, Zusammenspiel mit anderen Verarbeitungstätigkeiten |
| 42 | **8.E.** | Dokumentation der Maßnahmen zur Gewährleistung einer hohen Datenqualität und einer hohen Zuverlässigkeit, insbesondere Dokumentation von Tests und der Freigabe sowie Dokumentation der Faktoren, die von dem Chatbot – mutmaßlich – verwendet werden |

| | | |
|---|---|---|
| 43 | **8.F.** | Programmtechnische Unterlassung bzw. Schließung von Schnittstellen bei Verarbeitungsverfahren und Komponenten und Gewährleistung des Erhalts dieser Maßnahme durch Regelnde Maßgaben zum Verbot von Backdoors sowie qualitätssichernde Revisionen zur Compliance bei der Softwareentwicklung |
| 44 | **8.G.** | Verschlüsselung von gespeicherten oder transferierten Daten sowie Prozesse zur Verwaltung und zum Schutz der kryptografischen Informationen |
| 45 | **8.H.** | Festlegung und Umsetzung eines Löschkonzepts |
| 46 | **8.I.** | Schutz vor äußeren Einflüssen, insbesondere vor unberechtigtem Zugriff auf oder unberechtigter Veränderung der Daten oder Prozesse |
| 47 | **8.J.** | Festlegung und Kontrolle organisatorischer Abläufe, interner Regelungen und vertraglicher Verpflichtungen (Verpflichtung auf Datengeheimnis, Verschwiegenheitsvereinbarungen usw.) |
| 48 | **8.K.** | Erstellung und regelmäßige Überprüfung eines Notfallkonzepts zur Wiederherstellung einer Verarbeitungstätigkeit auf Basis der Redundanz von Hard- und Software sowie Infrastruktur |
| 49 | **8.L.** | Festlegung von Voreinstellungen und Daten/Parametern beim Training, die die Verarbeitung der personenbezogenen Daten auf das für den Verarbeitungszweck erforderliche Maß beschränken |
| 50 | **8.M.** | Regelungen zur Kontrolle von Prozessen zur Änderung von Verarbeitungstätigkeiten |
| 51 | **8.N.** | Maßnahmen für differenzierte Einwilligungs-, Rücknahme- sowie Widerspruchsmöglichkeiten und Schaffung notwendiger Datenfelder zur Umsetzung dieser Maßnahmen |
| 52 | **8.O.** | Gewährleistung der Möglichkeit, die Hauptanwendung zu nutzen, wenn der Chatbot nicht funktioniert, die Einwilligung verweigert oder widerrufen oder die Nutzung auf andere Weise abgelehnt wird |