KONZEPT FÜR DIE KLASSIFIKATION VON DIGITALEN DOKUMENTEN UND INFORMATIONEN DER STADTVERWALTUNG RAPPERSWIL-JONA

Bachelor Thesis

HWZ Hochschule für Wirtschaft Zürich

eingereicht bei:

Erwin Fang

Vorgelegt von:	Mauro De Cambio
Matrikelnummer:	20-678-942
Studiengang:	BWI-A20
Ort, Datum:	Jona, 22.05.2024

Abstract

Diese Thesis befasst sich mit der Entwicklung eines Konzepts zur Datenklassifizierung für die Stadtverwaltung Rapperswil-Jona. Ziel der Thesis ist es, die Datensicherheit zu verbessern und die Einhaltung des „Need to Know“-Prinzips durch die Einführung eines Datenklassifikations - und Data Leak Prevention-Systems zu unterstützen und zu verbessern. Das Konzept integriert vier Klassifikationsstufen mit jeweils unterschiedlichen Auswirkungen auf die entsprechend eingestufte Datei. Durch Data Leak Prevention-Richtlinien entsteht eine feingranulare und dynamische Steuerung von Datenzugriffen und Datenflüssen. Durch die Analyse der bestehenden Datenarchitektur und des Berechtigungskonzepts der Stadtverwaltung wurden Schwachstellen identifiziert und durch das vorgeschlagene Konzept adressiert.

Anhand der zuvor definierten Anforderungen wird überprüft, ob sich ein solches Konzept in der Stadtverwaltung umsetzen lässt. Der Proof of Concept zeigt auf, dass das entwickelte Berechtigungskonzept technisch und organisatorisch realisierbar und in vorhandene Systeme integrierbar ist. Er zeigt ausserdem auf, dass sich die Sicherheitslage durch ein Klassifikationssystem deutlich verbessern lässt. Herausforderungen, die während und nach der Implementierungsphase auftreten können, wurden ebenso beleuchtet wie der Bedarf an kontinuierlicher Schulung der Mitarbeitenden in den Bereichen Datensicherheit und Klassifikation. Die Ergebnisse dieser Arbeit verdeutlichen die Bedeutung einer sorgfältigen Datenklassifikation und strukturierter Sicherheitsrichtlinien, um die Sicherheit und Integrität sensibler Daten zu gewährleisten.

Mit dem entstandenen Konzept und dem Proof of Concept wird für die Stadtverwaltung Rapperswil-Jona ein wichtiger Meilenstein erreicht. Es bietet die Ausgangslage für die Realisierung der Datenklassifikation und der Einführung ausführlicher Data Leak Prevention-Massnahmen, um die Datensicherheit massgeblich zu verbessern und Datenlecks zu vermeiden.

Inhaltsverzeichnis

1 Einleitung

> "Profitability of organizations is ultimately dependent on the effectiveness with which they exchange, gather, process, retrieve, link, control, share, manage and, above all, protect their data and information. All these processes, however, require that the right information be made available to the right person or persons at the right place and at the right time."*(Eloff et al., 1996)*

Obwohl das Zitat bereits vor 28 Jahren veröffentlicht wurde, hat es bis jetzt nichts von seiner Aussagekraft verloren. Im Gegenteil, es trifft mehr denn je auf die heutigen Unternehmen und deren IT-Landschaften zu. Informationen sind in vielen Organisationen das wertvollste Gut. Wer gut mit Informationen umgehen kann, hat im heutigen Informationszeitalter einen klaren Vorteil.

Durch die digitale Transformation steht die Stadtverwaltung Rapperswil-Jona vor der Herausforderung, ihre wachsenden Datenmengen besser zu verwalten und zu schützen. Täglich wird durch die verschiedenen Abteilungen der Stadtverwaltung eine Vielzahl an unterschiedlichsten Daten generiert, verarbeitet und gespeichert. Diese reichen von sensiblen personenbezogenen Informationen, beispielswiese Steuerinformationen, bis hin zu Daten, die für die Planung und Entwicklung der Stadt Rapperswil-Jona und ihrer Infrastruktur von hoher Bedeutung sind.

Die unterschiedlichen Datentypen und Speicherorte, welche sich über die Zeit entwickelt haben, erfordern ein umfassendes Datenmanagement und -sicherheitskonzept. Bisher wurde der Zugang zu diesen Daten über ein Discretionary Access Control (DAC)-Berechtigungssystem gesteuert, welches zusammen mit Access Control Lists (ACLs) mit den ersten Microsoft-Computer in der Stadtverwaltung eingeführt wurde. Das ACL-System ermöglicht eine granulare Definition von Benutzerrechten auf Datei- und Verzeichnisebene. Trotz seiner eher einfachen Funktionsweise und weitverbreiteten Anwendung in der Informatik weist dieses System im Angesicht der neuen digitalen Herausforderungen der Stadtverwaltung bedeutende Limitationen auf.

Dokumente und Dateien befinden sich nicht mehr ausschliesslich auf lokalen Servern und Infrastrukturen. Eine Vielzahl von Daten wird mittlerweile in verschiedenen Cloud-Diensten gespeichert. Seit der Einführung von Microsoft 365 während der Corona-Pandemie 2020 verlagern sich Dokumente immer mehr in die beiden Microsoft-Dienste Microsoft OneDrive und Microsoft SharePoint. Durch die Anforderung der Kollaboration unter Mitarbeitenden aber auch mit externen Personen erfolgt vermehrt der Datenaustausch über E-Mail oder anderen kollaborative Tools

wie Microsoft Teams. Diese Verschiebung zu dezentralen und Cloud-basierten Speicherlösungen verlangt eine Neuausrichtung und Anpassung der Zugriffsmechanismen, um die Verfügbarkeit, Sicherheit und Integrität der Informationen der Stadtverwaltung Rapperswil-Jona zu gewährleisten.

1.1 Zielsetzung

Die Einführung einer Datenklassifikation stellt einen wichtigen Schritt in der Entwicklung eines Datenverwaltungskonzeptes für die Stadtverwaltung Rapperswil-Jona dar. Das Hauptziel dieses System ist es, die Sicherheit und Verwaltung von Daten zu verbessern, unabhängig davon, ob sie in dezentralen Cloudsystemen oder in der eigenen IT-Infrastruktur gespeichert sind. Die Klassifikation von Daten ist eine bewährte Praxis, die nicht nur in der Informatik Anwendung findet. Auch wenn die grundlegenden Prinzipien des Schutzes sensibler Informationen gleichbleiben, erfordern die modernen Informatiklandschaften eine neue und innovative Herangehensweise.

Das Konzept der Datenklassifikation ist grundsätzlich einfach. Durch die Analyse von Inhalten, wie beispielsweise personenbezogenen Daten wie Namen, Adressen oder AHV-Nummern, wird jedes Dokument einer bestimmten Sicherheitsstufe zugeordnet. Die Einteilung erfolgt in verschiedene Kategorien, die von der Organisationsstruktur und den spezifischen Sicherheitsanforderungen eines Unternehmens abhängen.

Im Zentrum dieser Thesis steht folgende Forschungsfrage:

Inwiefern trägt die Entwicklung und Implementierung eines Konzepts zur Datenklassifikation dazu bei, das Need-to-Know-Prinzip und die Data Loss Prevention der Stadt Rapperswil-Jona zu verbessern?

Somit ist das Ziel dieser Thesis, ein Konzept zur Nutzung der Datenklassifikation in der Stadtverwaltung Rapperswil-Jona zu entwerfen. Dieses Konzept soll die theoretischen Bedingungen der Datenklassifikation erläutern und durch ein Proof of Concept (PoC) die praktische Anwendbarkeit des Systems demonstrieren. Der PoC soll die Fähigkeit des Systems schaffen, Klassifizierungen vorzunehmen und Klassen zu vergeben, um das Risiko von Datenlecks zu minimieren.

Zusätzlich wird die Thesis die organisatorischen und technologischen Herausforderungen untersuchen, die mit der Einführung eines solchen Systems verbunden sind. Somit können Lösungen für eine Integration in bestehende und zukünftige IT-Strukturen der Stadtverwaltung vorschlagen werden. Die Thesis wird einen wichtigen Beitrag zur Sicherheitsstrategie der Stadtverwaltung

Rapperswil-Jona leisten, indem sie eine solide Basis für den Schutz und die effektive Verwaltung der Informationen der Stadtverwaltung legt.

1.2 Aufbau der Arbeit und methodische Vorgehensweise

Zu Beginn der Thesis steht eine umfassende Literaturrecherche, welche darauf abzielt, einen Überblick über den aktuellen Stand der Forschung im Bereich der Datenklassifikation zu gewinnen. Dieser Teil beinhaltet die systematische Suche nach für die Beantwortung der Forschungsfrage relevanten wissenschaftlichen Büchern, Artikeln und anderen Veröffentlichungen. Es wird vertieft auf die bestehenden Methoden zur Datenklassifikation eingegangen und grundlegende Konzepte dazu erläutert. Die Ergebnisse der Literaturrecherche dienen dazu, eine theoretische Grundlage für die Entwicklung des Klassifikationskonzepts für die Stadtverwaltung Rapperswil-Jona zu schaffen.

Basierend auf den Erkenntnissen aus der Literaturrecherche wird ein Konzept zur Datenklassifikation entwickelt. Dies beinhaltet das Definieren von Methoden und Kriterien, welche zur Klassifikation von Daten verwendet werden sollen. Es werden Schwachstellen im aktuellen Sicherheitskonzept aufgezeigt und analysiert. Das Konzept wird technologieunabhängig erstellt und es werden technische sowie organisatorische Anforderungen für eine einheitliche Datenklassifikationslösung erarbeitet.

Im letzten Teil der Thesis wird das entwickelte Konzept anhand eines Proof of Concept auf die Implementierbarkeit überprüft. Dabei wird das Konzept anhand einer Softwarelösung testweise in kleinem Rahmen implementiert, um die Funktionalität und Effektivität zu testen. Diese Phase beinhaltet die Implementierung einer Klassifikationslösung und die Analyse der Ergebnisse. Der PoC bildet die Ausgangslage für die zukünftige Realisierung einer umfangreichen Datenklassifikationslösung für die Stadtverwaltung Rapperswil-Jona.

2 Grundlagen und Begriffsdefinition

Um die Komplexität von modernen Datenverwaltungs- und Sicherheitssystemen zu verstehen, ist es wichtig, ein gemeinsames Verständnis der grundlegenden Prinzipien und der verwendeten Begriffe zu etablieren. Dieses Kapitel bildet die Grundlage für die Erarbeitung des Klassifikationskonzeptes, indem Konzepte und Fachausdrücke definiert werden.

2.1 Das Need-to-Know-Prinzip

Das «Need-to-Know»-Prinzip (auch Least-Privilege-Prinzip) ist ein wichtiger Bestandteil der Informationssicherheit. Es befasst sich mit der Kontrolle des Verbreitens und des Zugriffs auf Informationen. Das Prinzip verlangt, dass Personen oder Systeme nur Zugriff zu Informationen erhalten sollen, welche für die Ausführung ihrer Arbeit oder Aufgabe von Nöten sind. Damit sollen sensible Information vor unberechtigtem Zugriff geschützt werden. In Unternehmen wird das «Need-to-Know»-Prinzip häufig eingesetzt, um das Risiko von Datenlecks, das heisst den Abfluss von Daten aus dem Unternehmen, zu verhindern.

Beim «Need-to-Know»-Prinzip wird die Ansicht vertreten, dass nicht jede Person innerhalb eines Unternehmens Zugang zu allen Informationen benötigt, um ihre Arbeit auszuführen. Doch diese Ansicht beschränkt sich nicht nur auf Menschen. Auch Systeme sollen nur auf die Informationen zugreifen können, welche für die ordentliche Funktionstüchtigkeit nötig sind.

Für die Umsetzung des «Need-to-Know»-Prinzips wird in Unternehmen eine Kombination aus technischen Systemen und organisatorischen Massnahmen eingesetzt. Eine grosse Herausforderung bei der Implementierung des Prinzips stellt die Balance zwischen Datensicherheit und Zugänglichkeit dar. Grundsätzlich muss sichergestellt werden, dass Mitarbeitende nicht durch restriktive Massnahmen an ihrer Arbeit gehindert werden. Es ist jedoch kritisch, dass keine Risiken durch zu offene Zugriffsregeln entstehen. Zudem sind die Aufgaben und Rollen innerhalb eines Unternehmens nicht statisch. Die Zugriffsrechte erfordern somit dynamische Anpassungen, wenn sich das Unternehmen weiterentwickelt.

2.2 Zugriffsmodelle

Um das «Need-to-Know»-Prinzip erfolgreich umzusetzen, kann auf verschiedene Zugriffsmodelle rekurriert werden. In diesem Kapitel wird vor allem auf zwei Modelle vertieft eingegangen: Die «Discretionary Access Control» (DAC) und die «Role Based Access Control» (RBAC). DAC ist das das Zugriffsmodell, welches zum jetzigen Zeitpunkt bei der Stadtverwaltung eingesetzt wird. Es

ermöglicht feingranulare Kontrollen von Berechtigungen, welche jedoch angesichts der zunehmenden Komplexität und Dezentralisierung das «Need-to-Know»-Prinzip nichtmehr ausreichend erfüllen. Das Modell wird in dieser Thesis auf seine Eignung hin kritisch reflektiert und analysiert.

Als zweites Modell wird RBAC erläutert. Das Modell soll zukünftig bei der Stadtverwaltung zum Einsatz kommen. RBAC verfolgt den Ansatz, Berechtigungen nicht mehr wie bei DAC einem Benutzeraccount selbst zuzuteilen, sondern Zugriffe auf Basis von Rollen zu verwalten. In Verbindung mit der Datenklassifikation, welche im Folgenden noch weiter thematisiert wird, verspricht RBAC eine bessere Übereinstimmung mit den Datensicherheitszielen der Stadtverwaltung.

2.3 Discretionary Access Control

Bei «Discretionary Access Control» (DAC)-Modellen werden die Zugriffs-Aktionen auf Ressourcen von spezifischen Benutzern explizit in Regeln definiert (Petkovii & Jonker, 2007). Jeder Ressource wird demnach für jede zugreifende Instanz (User oder System) eine Berechtigung zugeteilt. In einer DAC-Umgebung können die Besitzer einer Ressource selbst die Berechtigung auf diese Ressource steuern. Den Besitzern ist also erlaubt, Benutzerberechtigungen anzupassen, zu entfernen oder neuen Benutzern den Zugriff auf die Ressource zu gewähren.

Beim Zugriff auf eine Ressource wird überprüft, ob der zugreifende User oder das zugreifende System die für die Aktion erforderlichen Berechtigungen besitzt. Für diese Überprüfung wird die Identität der zugreifenden Instanz überprüft. Wir nehmen an, dass **Jon** (zugreifende Instanz) die Datei **Bewerbung** (Ressource) aufrufen und **lesen** (=read – Berechtigung) will. In diesem Fall wird überprüft, ob der Benutzer **Jon** (oder eine Gruppe, in der Jon Mitglied ist) die Berechtigung **read** auf das Objekt **Bewerbung** besitzt. Falls ja, wird ihm der Zugriff gewährt und der Benutzer Jon darf die Datei lesen. Implementiert wird DAC anhand von Access Control Lists (ACLs).

2.3.1 Access Control Lists

ACLs (Access Control Lists) bestimmen in einer DAC-Umgebung, wer auf welche Dateien zugreifen darf. Bei der zugreifenden Instanz kann es sich um einzelne Benutzende, Benutzergruppen oder um Systemprozesse handeln. ACLs sind in der IT kein neues Konzept. Bereits früh wurden ACLs von Computersystemen genutzt, um den Zugriff von Usern auf Ressourcen innerhalb eines Systems zu kontrollieren. So erschien in der Zeitschrift «Computerworld» vom 21. Mai 1984 auf Seite 54 ein Artikel mit folgendem Wortlaut:

« The new Version of Gent-II (revision 3.0) has added a line-security mechanism which is implemented under the Primos ACL subsystem. In a networking environment, the software product is said to provide a means to reserve lines, to give certain users privileged access to all lines and to forbid use of specified lines .» (Enterprise, 1984)

Eine ACL beinhaltet keine oder mehr Access Control Entries (ACE). Diese ACEs verbinden eine Identität mit einer Berechtigung. Somit kann festgelegt werden, wie die entsprechende Identität mit der Ressource interagieren darf (Ashcraft et al., 2021, 2023). So wird nicht nur der Zugriff, sondern auch die Operation (lesen, schreiben, ausführen), welche mit der Ressource durgeführt werden darf, geregelt.

ACLs ermöglichen so Administratoren die granulare Kontrolle über kritische Informationen. In der Praxis führt jedoch genau diese feine Kontrolle über Zugriffe zu immer mehr Herausforderungen. Durch die stetig wachsenden Datenmengen in Organisationen führt die mangelnde Skalierbarkeit von ACLs in komplexen Umgebungen zu enormem Arbeitsaufwand. Dateien und Ordner in Umgebungen mit mehreren hunderttausenden Objekten einzeln und von Hand zu berechtigen ist schlicht nicht praktikabel. ACLs wirken dem mit den beiden Funktionen Berechtigungsgruppen und Vererbung entgegen.

2.3.2 Berechtigungsgruppen

Die gängigsten Betriebssysteme bieten die Möglichkeit, Benutzergruppen für einzelne Ressourcen zu berechtigen. Benutzergruppen bestehen aus einzelnen Benutzern. Diesen Benutzergruppen können dann entsprechende Berechtigungen für die Ressourcen zugewiesen werden. Alle Benutzer innerhalb dieser Benutzergruppe erhalten die gleichen Berechtigungen für die Ressource wie die Benutzergruppe selbst. So müssen keine Berechtigungen für einzelne Benutzer gesetzt werden. Dies vereinfacht die Handhabung stark, da die Benutzergruppen meist zentral in einem Benutzerverzeichnisdienst verwaltet werden.

2.3.3 Vererbung

Auch wenn die Möglichkeit, Benutzergruppen auf Ressourcen zu berechtigen, die Komplexität verringert, ist der Aufwand hoch. Aus diesem Grund wird die Vererbung eingesetzt. Bei der Vererbung werden die Berechtigungen, welche auf die übergeordnete Ressource angewendet wurden, weitergegeben. So können beispielsweise Dateien innerhalb eines Ordners von Benutzern gelesen werden, auch wenn diese nur für den übergeordneten Ordner explizit berechtigt wurden. Somit müssen nicht mehr für jede einzelne Ressource Berechtigungen gesetzt werden. Dies hat

den Vorteil, dass nur die in der Dateistruktur am höchsten angeordnete Elemente verwaltet werden müssen. Ein wesentlicher Nachteil bei der Vererbung ist jedoch, dass keine feineren Berechtigungen innerhalb der Struktur angewendet werden können.

Trotz Berechtigungsgruppen und Vererbung werden Berechtigungsstukturen mit ACLs in grösseren Unternehmen zunehmend unübersichtlich und fehleranfällig. Die statische Zugriffskontrolle erschwert zudem die Implementierung von ACLs in modernen dynamischen Cloud-Umgebungen.

2.4 Role Based Access Control

Role Based Access Control ist ein Ansatz, bei welchem die Zugriffsrechte nicht den einzelnen Benutzern, sondern Rollen zugewiesen werden (Osborn et al., 2000). Dies steht im Vergleich zu dem im vorherigen Kapitel beschriebenen ACLs, bei welchen die Zugriffsrechte direkt der Identität zugewiesen sind. Dies soll dazu dienen, das Management von Berechtigungen in komplexen Datenumgebungen zu vereinfachen. Rollen und Gruppen unterscheiden sich darin, dass eine Gruppe eine Ansammlung von Benutzern beinhaltet, welche die gleichen Berechtigungen, die der Gruppe zugewiesen sind, beinhaltet. Eine Rolle wiederum ist eine Sammlung von spezifischen Berechtigungen, welche einem Benutzer zugewiesen werden. Rollen sind also benutzerunabhängig. Sie werden meist anhand der Rollen der Mitarbeitenden im Unternehmen definiert. Eine Rolle kann zum Beispiel «Administration Informatikdienst» sein. Wenn diese Rolle auf die Dateiablage «Administratives» Lese- und Schreibrechte besitzt, erhalten die Benutzer, welchen diese Rolle zugewiesen ist, ebenfalls diese Lese- und Schreibrechte. Solche Rollen können also als Ansammlung von Berechtigungen auf gewisse Ressourcen angesehen werden.

Aus dem Konzept der Rollen innerhalb von RBAC ergeben sich einige Vorteile gegenüber DAC. Für Administratoren ist es einfacher, vordefinierte Rollen an Benutzer zuzuweisen als einzelne Berechtigungen zu setzen. Durch Rollen ist es zudem deutlich einfacher nachzuvollziehen, welche Zugriffsrechte für welche Benutzer im Einsatz sind (Osborn et al., 2000). So können in komplexeren Umgebungen mit bedeutend weniger Zeitaufwand Berechtigungen verwaltet und überwacht werden. Ein weiter Vorteil von Rollen ist die Skalierbarkeit. So können Wachstum oder Änderungen der Unternehmensstruktur effizient mit Rollen abgebildet werden.

2.5 Attribute Based Access Control

In der NIST-Publikation 800-162 im Januar 2014 wird beschrieben, dass ACLs und RBAC als spezielle Anwendungen von Attribute Based Access Control (ABAC) gesehen werden können (Hu et

al., 2014). Dabei nutzen ACLs das Attribut der Identität und RBAC das der Rolle. Im Gegensatz dazu stützt sich ABAC auf ein System aus Richtlinien, welche die Bewertung einer Vielzahl verschiedener Attribute ermöglichen. Attribute sind Eigenschaften, welche die verschiedenen Teilnehmer eines Requests mitbringen. Sie beinhalten Informationen, welche entscheiden, ob ein Request genehmigt wird oder nicht.

Teilnehmer eines Requests können sein (Hu et al., 2014):

- Benutzer, Services oder Geräte, die einen Zugriff auf Ressourcen anfordern, um bestimmte Operationen durchzuführen.
- Ressourcen, die angefordert werden. Dazu können Dateien, Programme, Prozesse, Computer, Netzwerke und alles, worauf Operationen ausgeführt werden können, gehören.
- Operationen wie Lesen, Schreiben, Ausführen, Bearbeiten, Löschen und feinere Einschränkungen, zum Beispiel das Kopieren von Inhalten, das Ausdrucken von Dokumenten oder das Versenden per E-Mail oder Nachricht.
- Umweltzustände, die systemunabhängig auftreten, wie Datum und Zeit oder der Aufenthaltsort eines Benutzers.

In Richtlinien wird festgelegt, welche Attribute erfüllt sein müssen, um den geforderten Zugriff und die Ausführung der Operation zu gestatten.

Ein typischer Anwendungsfall von ABAC ist Geo-Blocking, das heisst das Blockieren oder Erlauben von Zugriffen auf ein System in Abhängigkeit von bestimmten geografischen Eigenschaften. Die Stadtverwaltung setzt bereits ein solches System ein. So wird der Zugriff auf die gesamte IT-Infrastruktur nur aus der Schweiz und aus umliegenden Ländern erlaubt. Versucht ein User von einem Gerät in der Schweiz auf eine Datei zuzugreifen, wird dies vom System gestattet. Versucht der User unter sonst gleichbleibenden Bedingungen den Zugriff aus China, wird diese Anfrage automatisch blockiert. Das verwendete Attribut ist in diesem Fall die ausgehende Ortschaft der Anfrage. Attribute können auch frei kombiniert werden. So könnte die Policy erweitert werden, um den Zugang aus China zuzulassen, sofern der Zugriff zwischen 10 und14 Uhr erfolgt (Hu et al., 2014).

2.6 Data Leak Prevention

Data Leak Prevention (auch Data Loss Prevention oder DLP) dient dazu, den ungewollten Abfluss sensibler Daten zu verhindern. Zusätzlich soll die Nutzung und die Verbreitung dieser Daten identifiziert, kontrolliert und überwacht werden (Föck & Fröschle, 2011).

So schnell die Digitalisierung von Informationen voranschreitet, so zahlreich sind die Risiken, dass eben diese Informationen in die Hände von nicht berechtigten Dritten gelangen. Die Risiken beziehen sich nicht nur auf externe Angriffe, sondern auch auf interne Bedrohungen. Informationen können auf unterschiedliche Weisen, beabsichtigt und unbeabsichtigt, aus dem Unternehmen abfliessen. So können Daten über Web-Applikationen (wie DropBox oder Texteditoren), über E-Mail oder Instant Messaging (wie Teams, WhatsApp oder private Mail Accounts), über externe Speichermedien oder per Druck von Dokumenten in die falschen Hände geraten (Raman et al., 2011).

DLP beinhaltet unterschiedliche Mechanismen und Strategien, um sicherzustellen, dass sensible Daten innerhalb festgelegter Grenzen eines Unternehmens verbleiben (Raman et al., 2011). So können unterschiedliche Tools wie Endpoint Security-Lösungen als Kontrolle auf einem Endgerät, Firewalls zur Kontrolle und Überwachung des Netzwerkes oder Tools zur Überwachung von Datenströmen und –interaktionen zum Einsatz kommen. Auch die Konzepte von RBAC, ABAC und DAC können als Teil einer DLP-Strategie angesehen werden. Sie bilden die erste Sicherheitsinstanz gegen unberechtigte Zugriffe auf Informationen (Raman et al., 2011).

Um eine effektive DLP-Strategie anwenden zu können, ist es wichtig, dass ein Unternehmen weiss, welche Informationen es besitzt (Wlosinski, 2018). Nur so kann sichergestellt werden, dass ein Unternehmen auch weiss, welche Informationen sensibel und somit schützenswert sind. Es sollte definiert werden, welche Informationen welchen Wert im Unternehmen besitzen und mit welchem Schaden gerechnet werden muss, sollten diese Daten an unberechtigte Dritte gelangen. Es ist wichtig, dass die Herkunft, der Wert und der Speicherort der Daten identifiziert werden. Sind länderspezifische Anforderungen an den Datenschutz und die Datensicherheit vorhanden, müssen diese eingehalten werden.

Eine Analyse der bestehenden Sicherheitsinfrastruktur, der Sicherheitslücken und Risiken bildet die Grundlage, um notwendige Schritte zum Schutz der Daten einzuleiten (Wlosinski, 2018). So kann früh erkannt werden, wo eventuelle Anpassungen bei Hard- und Software nötig sind.

Eine klar definierte DLP-Strategie sollte die Prävention vor beabsichtigter und unbeabsichtigter Offenlegung von Daten, die Balance zwischen Datensicherheit und Datennutzbarkeit (Need to Know), den Schutz von Kundeninformationen sowie den Schutz von geistigem Eigentum und persönlichen Daten umfassen (Wlosinski, 2018).

Ein Risikoassessment sollte durchgeführt werden, um wirksame Richtlinien zu erstellen, welche der effektiven Umsetzung und Überwachung von DLP dienen (Wlosinski, 2018). Diese sollten für

den internen Datenzugriff und Gebrauch und für den ausgehenden Datenfluss definiert werden. So können nicht nur Daten innerhalb des Unternehmens geschützt werden, sondern es kann auch geklärt werden, wie mit dem Versand von Daten an Dritte umgegangen wird.

Ein grosser Bestandteil von DLP sind die Mitarbeitenden. Es ist wichtig, dass das unternehmensweite Bewusstsein für die Datensicherheit gestärkt wird (Wlosinski, 2018). Dies kann durch den Einsatz von Schulungen oder Workshops geschehen, in denen die Mitarbeitenden über Datenschutz und Datensicherheit aufgeklärt und sensibilisiert werden. So kann beispielsweise die korrekte Nutzung von E-Mail, Internet und Verschlüsselung informiert werden.

Wie bereits erwähnt beinhaltet eine DLP-Strategie verschiedene Tools, Technologien und Mechanismen, um Datenverlust zu vermeiden. Ein Konzept, welches in dieser Thesis vertieft betrachtet wird, ist, Daten zu labeln und zu klassifizieren. Die folgenden Kapitel erläutern die grundlegende Funktionsweise dieses Konzeptes.

2.6.1 Labeling und Klassifizierung von Daten

Daten werden in der modernen Arbeitswelt nicht mehr nur innerhalb des firmeneigenen Netzwerks gespeichert. Durch weltweit verfügbare Cloud-Infrastrukturen, welche direkt über das Internet erreichbar sind, sind Daten jederzeit auf unterschiedlichsten Endgeräten aufrufbar. Eine Einschränkung der bisher beschriebenen Sicherheitskonzepte (RBAC, ABAC, DAC, ACL) ist, dass diese Dateien nur schützen, solange diese innerhalb von vom Unternehmen verwalteten Infrastrukturen vorhanden sind. Wird beispielsweise eine Datei von einem autorisierten User auf ein privates Endgerät heruntergeladen, so hat das Unternehmen ab diesem Zeitpunkt keinerlei Kontrolle mehr über diese Daten. Dieser Problematik sollen DLP und Datenklassifikation entgegenwirken.

Ein weiterer wichtiger Aspekt der DLP ist, dass ein Unternehmen die eigenen Daten kennt. Nur wenn bekannt ist, wie sensibel Daten sind, können diese entsprechend geschützt werden. Um diesen Prozess des Kennens und Schützens möglichst effizient und einfach zu gestalten, können Dokumente aufgrund von Eigenschaften in Kategorien, sogenannte Klassen, eingeteilt werden. Unterschiedliche Klassen werden für verschiedene Sensibilitäten verwendet. Aufgrund von Sicherheitsrichtlinien werden Dateien dann auf Grund ihrer Klassifikation von Sicherheitssystemen, Applikationen und Mitarbeitenden unterschiedlich behandelt. Durch die Klassifizierung können schnell Regeln auf grosse Mengen an Informationen angewendet werden.

Die Klassifizierung erfolgt anhand sogenannter Label (Bailey, Robertson, et al., 2024). Einer Datei wird ein Label entweder manuell oder durch verschiedene automatische Methoden zugewiesen. Ein Sicherheitssystem wendet dann aufgrund des zugeteilten Labels Regeln auf diese Datei an. Diese Regeln beinhalten beispielsweise auch die Zuweisung in eine Klasse. Wird einem Dokument ein Label zugwiesen, bleibt dieses an diesem Dokument haften. Das Label wird in die Metadaten der Datei geschrieben. Wird das Dokument beispielsweise per Mail verschickt, kopiert oder verschoben und an einem anderen Ort gespeichert, so ist es auch da noch mit diesem Label gekennzeichnet. So wird sichergestellt, dass ein Dokument unabhängig vom Endgerät und vom Speicherort immer das zugewiesene Label besitzt.

Durch die Label können Dokumente einer vorgegeben Sicherheitsrichtlinie zugwiesen werden. Zum Beispiel können Dokumente, welchen das Label «geheim» zugwiesen sind, automatisch verschlüsselt werden. Endanwendungen, welche mit dem Labeling-System kompatibel sind, können ebenfalls aufgrund des Labels gewisse Aktionen erlauben oder verbieten. So kann beispielsweise ein Mailprogramm verhindern, dass eine Datei mit dem Label Vertraulich als Anhang an eine externe Empfängeradresse versendet wird.

Eine Möglichkeit, Dokumenten ein Label bzw. eine Klassifikation zuzuweisen, besteht darin, der Person, welche das Dokument erstellt, die Auswahl eines Labels zu überlassen. Bei gängiger Klassifizierungssoftware wird beim Speichern der Datei eine Auswahl mit den Klassen angezeigt. Wird eine dieser Klassen ausgewählt, so erhält das Dokument das entsprechende Label und wird der Klasse zugeteilt. Wird kein Label ausgewählt, wird das Dokument automatisch mit einem Standard-Label versehen und der Standardklasse zugeteilt. Dies könnte beispielsweise die Klasse «Vertraulich» sein. Dieser Vorgang setzt die Kompatibilität der Applikation mit der Klassifikationslösung voraus und kann sich auf Grund der Applikation unterscheiden. Die manuelle Klassifikation durch die Mitarbeitenden selbst setzt voraus, dass diese wissen, wie der Informationsgehalt in den von Ihnen erstellten oder bearbeiteten Dateien eingestuft werden muss.

In einem Unternehmen sind bei der Einführung einer DLP-Strategie meist bereits grosse Mengen an Daten vorhanden. Entsprechend hoch ist der Aufwand, alle vorhandenen Daten mit Labels zu versehen. So bieten die meisten gängigen DLP-Softwarelösungen eine Möglichkeit, bestehende und neue Daten automatisch mit Labels zu versehen. Es gibt unterschiedliche Ansätze, wie die Sensitivität von Daten erkannt werden kann und diese können sich je nach DLP-Lösung unterscheiden. Folgend sind einige gängige Möglichkeiten zum automatischen Erkennen von Text erläutert.

Exact Data Match

Beim Exact Data Match (EDM) handelt es sich um das Erkennen genauer Übereinstimmungen innerhalb von Datensätzen. EDM funktioniert, indem es Hashwerte innerhalb einer Datenbank mit den Daten abgleicht (Fox, Koenen, & Savell, 2024). Dabei wird nicht der Strings selbst, sondern die Hashwerte der Information in der Datenbank und im Datensatz verglichen (Exact Data Matching (EDM), 2024). Die Datenbank kann nicht nur die genaue Information, das primäre Element, enthalten, sondern auch unterstützende Elemente (Supporting Elements) (Fox, Koenen, & Savell, 2024). Diese unterstützenden Elemente ermöglichen eine feinere Kontrolle der Erkennung der primären Elemente. So kann festgelegt werden, dass das primäre Element im Datensatz nur mit der Datenbank übereinstimmt, wenn unterstützende Elemente innerhalb eines vordefinierten Abstands zum primären Element vorkommen. Beispielsweise soll eine spezifische AHV-Nummer (Schweizerische Sozialversicherungsnummer) innerhalb der Dateiablage erkannt und eingestuft werden. In der Datenbank wird also genau diese AHV-Nummer eingetragen. Zusätzlich zur AHV-Nummer werden für die AHV-Nummer unterstützende Elemente erfasst. Dies können beispielsweise der Vorname und Nachname der Person mit dieser AHV-Nummer sein. Der Datensatz in der Datenbank würde also folgendermassen aussehen:

AHV-Nummer **(Primary Element)**	**Name** **(Supporting Element)**	**Vorname** **(Supporting Element)**
756.5862.6697.14	Jon	Doe

Tabelle 1 Beispiel einer EDM-Tabelle

Wird nun in einer Datei genau diese AHV-Nummer erkannt, wird zusätzlich nach den unterstützenden Elementen gesucht. Kommt also in dem Dokument zusätzlich der Name Jon Doe vor, wird eine Übereinstimmung erkannt und eine definierte Regel kann angewendet werden.

Pattern Matching

Pattern Matching funktioniert ähnlich wie Exact Data Match. Beim Pattern Matching werden jedoch nicht nach genauen Informationen gesucht, sondern es wird anhand von Mustern überprüft, ob es zu einer Übereinstimmung kommt (Fox & Koenen, 2024b). Für Pattern Matching werden üblicherweise Regular Expressions, kurz Regex, eingesetzt. Mithilfe von Regex können innerhalb von Texten Zeichenfolgen anhand von Mustern gesucht werden. Durch die Flexibilität von Regex können Muster von simplen Ansammlungen von Buchstaben bis hin zu komplexen Zeichenfolgen abgedeckt werden (Friedl, 2006). Wenn wir beim Beispiel der AHV-Nummer bleiben, wird diese beim Pattern Matching nicht mehr gegen den Hash-Wert des genauen Strings getestet, sondern gegen die Regular Expression. Diese könnte im Falle der AHV-Nummer so aussehen:

```
756\.?[0-9]{4}\.?[0-9]{4}\.?[0-9]{2}
```

Anhand dieser Regular Expression (Regex) können nun alle AHV-Nummern erkannt werden. Bei EDM wären es nur die AHV-Nummern, welche auch in der Datenbank eingetragen sind.

Auch Pattern Matching kann mit unterstützenden Elementen verwendet werden, um die Genauigkeit der Übereinstimmung zu verbessern (Fox & Koenen, 2024b). So kann bei der AHV-Nummer festgelegt werden, dass innerhalb des festgelegten Bereiches die Wörter AHV oder Sozialversicherungsnummer vorkommen müssen. Damit sollen fälschlicherweise erkannte Übereinstimmungen (False Positives) verringert werden.

Document Fingerprinting

Das Document Fingerprinting wird eingesetzt, wenn standardisierte Dokumente geschützt werden sollen. Document Fingerprinting funktioniert, indem von einer leeren Vorlage eines Dokuments ein sogenannter Fingerabdruck erstellt wird (Fox & Koenen, 2024a). Beim Fingerabdruck handelt es sich üblicherweise um einen eindeutigen Hashwert. Die leere Vorlage, zum Beispiel ein nicht ausgefülltes Formular, wird analysiert, um Textmuster zu erkennen. So wird ein eindeutiger Hash des Dokumentes erstellt. Da beim Ausfüllen der leeren Vorlage nur Text hinzugefügt, nicht aber das ursprüngliche Textmuster verändert wird, kann die DLP-Software das ausgefüllte Dokument erkennen und DLP-Regeln darauf anwenden.

Document Fingerprinting funktioniert nicht bei Files, welche Password geschützt sind, ausschliesslich Bilder enthalten oder bei Dokumenten, welche den originalen Text der leeren Vorlage nicht mehr enthalten (Fox & Koenen, 2024a).

Machine Learning

Das rapide Wachstum der Mengen an Daten in der Stadtverwaltung macht es schwierig, alle Daten entsprechend ihrer Sensitivität richtig einzustufen und zu schützen. Viele moderne Datenklassifikations-Lösungen beinhalten die Möglichkeit, anhand von Machine Learning Algorithmen-Daten automatisiert zu klassifizieren. So können grosse Datenbestände effizient analysiert und eingestuft werden. Gängige Algorithmen zur Klassifikation von Text sind (Fong, 2010):

- Naive Bayes Classifier

Der Naive Bayes Classifier überprüft das Vorhanden- oder Abwesend-Sein von Merkmalen unabhängig von anderen Merkmalen (Ting et al., 2011). Merkmale sind im Falle der Textklassifizierung einzelne Wörter. Es wird überprüft, wie hoch die Wahrscheinlichkeit ist, dass ein Text zu einer Klasse gehört, wenn bestimmte Wörter vorhanden sind.

- Nearest Neighbor Classifier

Beim Nearest Neighbor Classifier (NNC oder auch KNN) wird ein Text mit anderen Texten verglichen und gleich wie in Bezug auf die Wortwahl und die Häufigkeit ähnlichsten Texte eingestuft (Fong, 2010). Texte werden als Vektoren dargestellt und die Ähnlichkeit zueinander berechnet. Der neue Text wird dann der Klasse seiner nächsten Nachbarn (k nächste Nachbarn) zugeordnet, wobei die Mehrheit entscheidet. Der Algorithmus ist einfach und flexibel, aber rechenintensiv und speicheraufwendig.

- Centroid based Classifier

Die Idee hinter dem Centroid based Classifier ist, dass sich alle Dokumente in einer Klasse einen einzelnen Repräsentanten (Centroid) teilen (Fong, 2010). Dieser Centroid kann beispielsweise der Mittelwert aller Texte in einer Klasse sein. Bei neuen Texten wird überprüft, zu welchem Centroid sie am besten passen, ähnlich wie beim Nearest Neighbor.

- Decision Tree

Decision Trees sind ein Ansatz zur Textklassifikation, der auf der Idee basiert, dass man eine Reihe von Entscheidungen treffen kann, um einen Text einer Klasse zuzuweisen (Fong, 2010). Dabei wird ein Baumdiagramm erstellt, welches an jeder Verzweigung eine Bedingung enthält, die auf den Merkmalen des Textes basiert. Der Pfad, dem man folgt, hängt davon ab, ob der Text die

Bedingung erfüllt oder nicht. Am Ende jedes Pfades befindet sich eine Klasse, die dem Text zugewiesen wird.

- Support Vector Machine

Bei der Textklassifikation mit Support Vector Machines (SVM) wird jedes Dokument in einem hochdimensionalen Raum dargestellt (Rigutini & Maggini, 2010). Jede Dimension stellt dabei ein Merkmal des Textes, wie zum Beispiel die Häufigkeit eines Wortes, dar. SVM verwenden dann eine lineare Hyperplane, um die Klassen optimal voneinander zu trennen. Neue Texte werden anschliessend klassifiziert, indem bestimmt wird, auf welcher Seite dieser Hyperplane sie liegen.

- Neuronal Network Classifier

Viele traditionelle Klassifikationsmethoden sind nicht in der Lage, den Kontext von Texten zu erkennen (Lai et al., 2015). Neuronale Netzwerke bieten die Möglichkeit, Texte nicht nur auf Grund der vorkommenden Wörter, sondern auch aufgrund des inhaltlichen Kontexts zu klassifizieren. Neuronale Netzwerke lernen, während des Trainings Merkmale aus den Texten zu extrahieren und diese so zu gewichten, dass sie der tatsächlichen Klasse der Trainingsdaten entsprechen. Nach dem Training können neue Texte klassifiziert werden, indem die gelernten Merkmale und Muster angewendet werden, um eine Vorhersage zu treffen.

3 Analyse der bestehenden Datenarchitektur und des Berechtigungskonzeptes

3.1 Struktur und Funktionsweise des aktuellen Berechtigungskonzeptes

Seit der Einführung von Windows-Computern setzt die Stadtverwaltung Rapperswil-Jona für die Zugriffskontrolle zu Dateien Access Control Lists (ACL) ein. Über die Zeit wuchsen die Datenmengen und somit auch die Anzahl Ordner und Dateien. Jeder Ordner und jede Datei benötigt entsprechende ACLs, um sicherzustellen, dass die zugreifende Instanz auch wirklich die Berechtigungen für die geforderte Operation besitzt. Dies führt dazu, dass die Komplexität der Berechtigungsstruktur stark anwächst.

Der Informatikdienst der Stadt Rapperswil Jona verwaltet insgesamt 6000 Benutzeraccounts in 2200 Benutzergruppen. Wenn für jeden Ordner und jede Datei einzeln manuelle Berechtigungen gesetzt werden, führt dies sehr schnell zu einem enormen Arbeitsaufwand und ist zudem äusserst fehleranfällig. So werden beispielsweise Berechtigungen von Mitarbeitenden, welche das Unternehmen verlassen haben, nicht automatisch entfernt oder bei einem Abteilungs- oder Funktionswechsel werden die Berechtigungen nicht entsprechend angepasst, was das Need-to-Know-Prinzip verletzt.

Zusätzlich zur wachsenden lokalen Infrastruktur werden Daten immer mehr in Cloud-Diensten verwendet. Die Stadt Rapperswil-Jona setzt seit Ende 2019 Microsoft 365 ein. Somit werden Daten nicht mehr nur in lokalen Verzeichnissen gespeichert, sondern finden ihren Weg zunehmend in verschiedenste Cloud-Applikationen.

3.2 Daten in der Cloud

Seit der Einführung von Microsoft 365 und dessen Services haben die Mitarbeitenden der Stadtverwaltung externer Leistungsbezieher Zugriff auf einen persönlichen Cloudspeicher bei Microsoft (OneDrive). Dieser Cloudspeicher ermöglicht es Benutzern, ihre eigenen Daten in der Microsoft-Cloudumgebung abzuspeichern und von beliebigen Geräten über einen Browser auf diese Daten zuzugreifen. Dies ermöglicht es, unabhängig vom Endgerät mit den eigenen Dokumenten zu arbeiten.

Daten in den OneDrive-Ablagen der Mitarbeitenden sind grundsätzlich nicht für andere Mitarbeite zugänglich. OneDrive bietet jedoch die Möglichkeit, dass Benutzer ihre Daten selbst mit

anderen Mitarbeitenden oder externen Personen teilen können. Dies wird durch die Funktion ermöglicht, Links für den Zugriff auf spezifisch definierte Daten zu generieren.

Noch ist ein Grossteil der Daten der Stadtverwaltung auf lokalen Speichern in den beiden eigenen Rechenzentren gespeichert. Es ist jedoch geplant, dass noch 2024 grossflächig Lokale Shares in die Microsoft Cloud (OneDrive und SharePoint) migriert werden. Obwohl diese Migration noch nicht gestartet wurde, steigt die Anzahl der Dateien und die Menge an verwendetem Onlinespeicher stetig an.

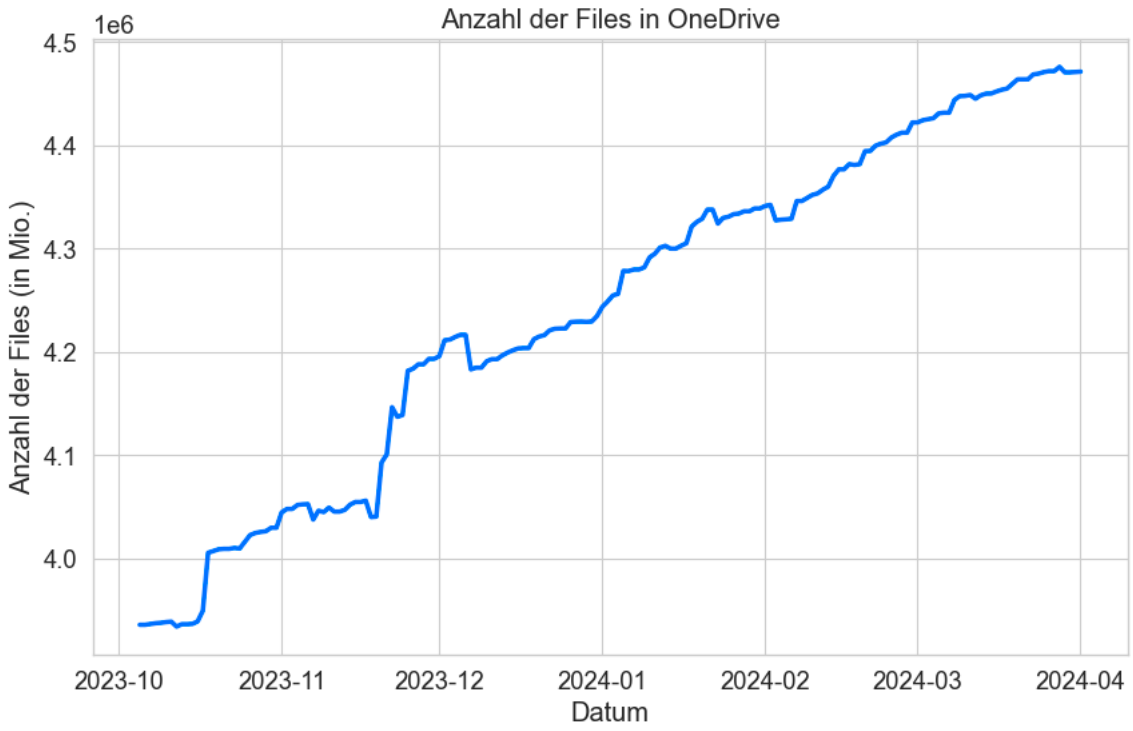

Abbildung 1: Anzahl Files in OneDrive

In Abbildung 1 ist zu sehen, wie die Datenmenge in OneDrive seit Oktober 2023 von 3.9 Millionen Files auf 4.4 Millionen Files Anfang April zugenommen hat. Dabei handelt es sich nur um Daten, welche Mitarbeitende selbst verwalten.

Zusätzlich zu OneDrive als persönlichem Speicherort wird SharePoint eingesetzt. SharePoint dient unter anderem als gemeinsam genutzte Dokumentenverwaltung. Obwohl auch die Datenmigration auf SharePoint noch bevorsteht, wächst die Datenmenge stetig an. Abbildung 2 ist zu entnehmen, dass die Anzahl der gespeicherten Dateien von knapp 650'000 Files im Oktober 2023 auf über 725'000 Files im April 2024 angewachsen ist. Die Daten wurden internen Berichten der Stadt Rapperswil-Jona über die Nutzung von OneDrive und SharePoint im Zeitraum von 180 Tagen entnommen.

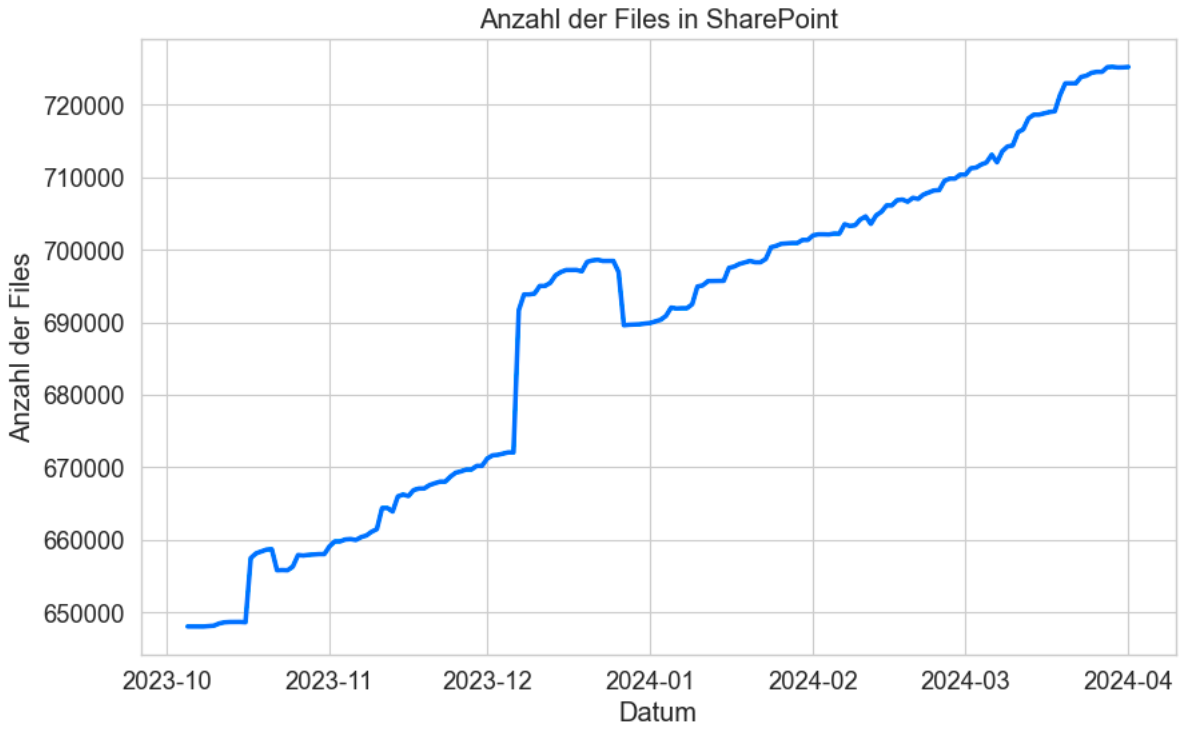

Abbildung 2: Anzahl Files in SharePoint

Auch SharePoint bietet die Möglichkeit, Daten mit internen und externen Personen zu teilen. Daten sind also nicht mehr nur vom intern verwalteten Netzwerk zugänglich, sondern können zunehmend über das Internet und von externen Personen aufgerufen werden. Das bedeutet, dass der Zugriff auf Dateien von unterschiedlichsten Endgeräten und Netzwerken geschieht. Mit der Möglichkeit, Links für den Zugriff auf Daten in der Cloudumgebung zu generieren, steigt das Risiko, dass ungewollt Zugriff auf heikle Daten gewährt wird.

3.3 Herausforderungen und Limitationen im aktuellen System

Die voranschreitende Verlagerung von Daten und Arbeitsprozessen in die Cloud, insbesondere in die Cloudumgebungen von Microsoft, bringt eine grosse Anzahl neuer Herausforderungen und Limitationen für das bestehende Berechtigungssystem mit sich. Bisher wurden Daten auf der stadteigenen lokalen Infrastruktur durch einfache Zugriffskontrollen (ACLs) geschützt.

Durch die verschiedenen Speicherorte (Cloud und lokal) ist es den Administratoren nicht mehr möglich, ACLs an einem zentralen Ort zu verwalten. ACL-Berechtigungen müssen zusätzlich in Cloud-Dateiablagen verwaltet werden. Dies führt zu weiterem Arbeitsaufwand und ist, wie auch die lokalen ACL-Berechtigungen, äusserst fehleranfällig.

Selbst mit Vererbung und Berechtigungsgruppen ist der Aufwand, eine solch komplexe Berechtigungsstruktur zu verwalten, sehr hoch. Durch die grosse Menge an Berechtigungen kann es schnell zu Fehlkonfigurationen kommen, welche ungewollten Zugriff auf sensible Informationen ermöglichen. Zudem ist es sehr schwer, ungewollte Zugriffe zu entdecken und zu beseitigen. Des Weiteren sind bestehende Berechtigungen auf statische Strukturen ausgelegt. Werden neue Berechtigungen verteilt, kann dies je nach Anzahl Ordner und Dateien Stunden dauern, bis die Berechtigungen auf alle Objekte vererbt wurden. Es werden wichtige Systemressourcen und vor allem viel Zeit benötigt. ACLs sind daher für den Gebrauch in einem dynamischen Arbeitsumfeld nur begrenzt einsetzbar. Ständig wechselnde Arbeitsrollen, Teams und Projekte benötigen Berechtigungsstukturen, welche sich mit dem Arbeitsumfeld verformen, anpassen und erweitern können. Zugriffe sollen sich nicht mehr nur auf die Identität des Benutzers beziehen, sondern auch den Kontext des Zugriffs (Standort, Gerätetyp usw.) berücksichtigen.

Durch die Verlagerung in die Cloud können Daten von unterschiedlichsten Endgeräten über das Internet abgerufen werden. Diese Geräte sind nicht nur Firmencomputer, sondern auch private Smartphones, Tablets, Notebooks und PCs. Diese Diversität erschwert die Verwaltung und Kontrolle der Berechtigungen enorm.

Eine Überwachung der Berechtigungen gibt es nur beschränkt. So werden Datenzugriffe und Änderungen aufgezeichnet, jedoch nicht, was nach einem Zugriff geschieht. So können User unbeaufsichtigt Informationen aus einem Dokument kopieren oder als Anhang per Mail verschicken. Es wird nicht aufgezeichnet, mit welchen Programmen ein User auf die Datei zugegriffen hat. Somit ist es nicht möglich, den Abfluss von Dateien in der Stadtverwaltung zu erkennen und frühzeitig zu verhindern.

In den neuen Arbeitsprozessen ist es nicht mehr ausreichend, nur den Zugriff auf eine Datei zu regeln. Informationen können über verschiedene Wege kopiert, geteilt, bearbeitet und über verschiedene Plattformen und Dienste verbreitet werden. Es ist also wichtig, die Information selbst unabhängig von ihrem Speicherort, oder der Form in der sie vorliegt, zu schützen.

4 Bedarfsanalyse für ein neues Berechtigungskonzept

Angesichts der identifizierten Limitationen des bestehenden Berechtigungssystems der Stadtverwaltung wird eine Bedarfsanalyse durchgeführt, um die Anforderungen an ein neues Berechtigungskonzept zu definieren.

4.1 Identifikation von Sicherheitsanforderungen

Klassifizierung neuer und vorhandener Daten

Bereits bestehende und neu erstellte Daten sollen **klassifiziert** werden. Es werden die vom Kanton St. Gallen empfohlenen Klassen verwendet (siehe Abschnitt Leitfaden des Kantons St. Gallen)**.** Dabei sollen Daten entsprechend ihrer enthaltenen Informationen in eine der vier Klassen **öffentlich**, **intern**, **vertraulich** und **geheim** eingeteilt werden. Aufgrund dieser Klassifikation müssen Regeln angewendet werden können.

Neu erstellte Dokumente erhalten standardmässig die Klassifikation **vertraulich**. Dadurch wird gewährleistet, dass Daten nicht versehentlich zu niedrig klassifiziert werde. Die Mitarbeiter werden dadurch angeregt, aktiv beim Erstellen einer Datei die Standardklassifikation zu hinterfragen.

Beim Erstellen oder Bearbeiten einer Datei soll aufgrund des Inhaltes automatisch eine Klassifikation vorgeschlagen werden. Gleichzeitig sollen den Mitarbeitenden die Gründe für die vorgeschlagene Klassifikation angezeigt werden. So verbessert sich das Verständnis der Mitarbeitenden bezüglich der Sensitivität von Informationen, die innerhalb einer Datei vorhanden sind.

Die erste Klasse, die einer Datei zugewiesen wird, dient als Ausgangslage für zukünftige Änderungen. So muss ein Wechsel der Klasse begründet werden. Dies geschieht entweder durch vorgefertigte Antworten oder eine benutzerdefinierte Eingabe. Der Wechsel der Klasse und die dazugehörige Begründung sollen aufgezeichnet werden.

Leitfaden des Kantons St. Gallen

Seit Anfang 2024 gibt es vom Kanton St. Gallen einen Leitfaden zur Klassifikation von Daten mit Microsoft 365. Die Stadtverwaltung Rapperswil-Jona ist bestrebt, diesem Leitfaden mit dem eigenen Sicherheitskonzept zu folgen.

Durch den flächendeckenden Einsatz von Microsoft 365 innerhalb des Kantons St. Gallen, wurde ein Leitfaden für die Klassifizierung von Daten erarbeitet. Ziel ist es, dass möglichst alle Gemeinden und Städte innerhalb des Kantons die Daten gleich klassifizieren. Somit soll sichergestellt werden, dass Daten untereinander ausgetauscht werden können, ohne die Klassifizierungsstufe ändern zu müssen. Ein Dokument mit der Klasse Vertraulich wird also von allen Städten und Gemeinden im Kanton St. Gallen als vertrauliches Dokument behandelt. Die Stadt Rapperswil-Jona als Teil des Kanton St. Gallen übernimmt diese Klassen. Der Kanton St. Gallen gibt folgende Klassen für die Datenklassifizierung vor:

Geheim	Vertraulich	Intern	Öffentlich

Folgend ist beschrieben, aufgrund welcher Eigenschaften ein Dokument mit einer Klasse versehen wird.

Geheim

Was	**Beschreibung**
Besonders schützenswerte Personaldaten (DSG, Art. 5)	Daten über: • Religion, Weltanschauung, politische oder gewerkschaftliche Ansichten oder Tätigkeiten • Gesundheitsdaten, Rasse oder Ethnie • Genetische Daten • Biometrische Daten • Strafrechtliche Verfolgungen • Massnahmen der sozialen Hilfe

Persönlichkeitsprofile oder Profiling (DSG, Art. 5)	Daten, um persönliche Aspekte einer natürlichen Person zu bewerten, analysieren oder vorherzusagen
Berufsgeheimnisse (StGB, Art 321)	Geheimnis, das infolge der Berufsausübung anvertraut worden ist.
Besondere Amtsgeheimnisse (OeffG, Art. 7)	Daten, deren Kenntnisnahme durch Unberechtigte öffentliche Interessen schwerwiegend beeinträchtigen kann
Vertraglich geschützte Daten	z.B. Geschäfts oder Fabrikationsgeheimnisse

Vertraulich

Was	**Beschreibung**
Allgemeines Amtsgeheimnis (OeffG, Art. 6 und 7)	Daten, deren Kenntnisnahme durch Unberechtigte öffentliche Interessen erheblich beeinträchtigen kann.

Intern

Was	Beschreibung
Informationszugang auf Anfrage (OeffG, Art. 5)	• Amtliche Dokumente, zu denen auf Anfrage Zugang gewährt werden muss. • Informationen zur Tätigkeit der Stadtverwaltung • Daten, deren Kenntnisnahme durch Unberechtigte öffentliche Interessen beeinträchtigen kann.

Öffentlich

Was	Beschreibung
Daten im Interesse der Allgemeinheit (Oeffg, Art 4)	Informationen über die Tätigkeiten der Stadtverwaltung, soweit diese von allgemeinem Interesse sind.
Öffentlich zugängliche Daten	Daten, die allen Personen zur Verfügung gestellt werden

4.2 Benutzeranforderungen

Die Ziele der Datensicherheit und Zugriffsteuerung stimmen nicht immer mit den Anforderungen der Benutzerfreundlichkeit überein. Umso wichtiger ist es, trotz strenger Sicherheitsvorschriften das Benutzererlebnis so angenehm wie möglich zu gestalten. Folgend sind Anforderungen definiert, welche dabei helfen sollen.

Klar definierte und dokumentierte Klassen

Klassen sollen klar definiert und dokumentiert sein. Mitarbeitenden muss stets klar sein, welche Klasse welche Auswirkungen auf die von ihnen verarbeiteten Daten hat.

Keine Einschränkung beim Daily Business

Durch die Klassifikation soll die alltägliche Arbeit so wenig wie möglich beeinflusst werden. Durch die Einhaltung des Need-to-Know-Prinzips sollen Daten immer durch die Mitarbeitenden verfügbar sein, wenn dies für das Ausüben der Arbeit nötig ist.

Einfache Anwendung

Die Klassen müssen einfach in den Applikationen, welche im Arbeitsalltag verwendet werden, integriert sein. Mitarbeitende sollen die Klassen ohne grossen zusätzlichen Aufwand einer Datei zuweisen können.

Schulungen

Mitarbeitende, die mit den Klassen arbeiten, müssen geschult werden. Schulungen sollen Wissen zur Klassifikation, deren Auswirkungen im Arbeitsumfeld und die korrekte Anwendung vermitteln. Zusätzlich ist die Sensibilisierung im Bereich Datenschutz und Datensicherheit ein wichtiger Aspekt.

4.3 Technische Anforderungen

Folgende technische Anforderungen sollten durch eine Software, welche für die Klassifikation eingesetzt wird, erfüllt werden.

- Die Software sollte in der Lage sein, Daten aus verschiedenen Quellen zu erfassen und zu klassifizieren. Quellen sind Systeme, welche von der Stadtverwaltung eingesetzt werden und Daten generieren und verarbeiten. Dazu zählen E-Mailsystem (Microsoft Exchange Online), Dateisysteme (Microsoft OneDrive und SharePoint), Cloudsysteme (Microsoft Azure) und Endpunkte (Windows-Endgeräte).
- Unterstützung einer Vielzahl von Datenformaten und -typen, einschliesslich strukturierter und unstrukturierter Daten, sowie die Fähigkeit, Daten über verschiedene Übertragungskanäle zu schützen.
- Es muss möglich sein, Datenklassen automatisch und manuell zuzuweisen. Da bereits grosse Datenmengen bestehen, muss das System in der Lage sein, sensible Daten automatisch zu erkennen und gemäss vordefinierter Kriterien zu klassifizieren. Bei neu erstellten oder bearbeiteten Dokumenten kann eine Klassifikation manuell vorgenommen werden.

- Detaillierte Sicherheitsrichtlinien müssen basierend auf der Datenklassifikation automatisch durchgesetzt werden können.
- Die Lösung sollte in der Lage sein, sensible Daten zu verschlüsseln und die Verwaltung von Zugriffsrechten zu unterstützen, um unbefugten Zugriff zu verhindern und so das Need-to-Know-Prinzip durchzusetzen
- Der Datenverkehr muss überwachbar sein und bei Verstössen gegen die Sicherheitsrichtlinien muss Alarm ausgelöst werden.
- Erfassung detaillierter Logs und Bereitstellung umfassender Berichte über Sicherheitsvorfälle, um Analysen zu ermöglichen und Compliance-Anforderungen zu erfüllen.
- Die Software muss in der Lage sein, Daten schnell zu erfassen und zu verarbeiten. Der Inhalt von neu erstellten Daten soll in Echtzeit erkannt werden.
- Das System soll in der Lage sein, mit dem zunehmenden Wachstum der Datenmengen zu skalieren
- Die Verwaltung der Klassifizierung und der Sicherheitsrichtlinien soll über eine Benutzeroberfläche (GUI) möglich sein.
- Berichte und Statistiken über Vorfälle, Systemstatus, verwendete Klassen, Datensensitivität können erstellt werden.
- Regelverstösse sollen automatische Aktionen hervorrufen, um Datenverlust frühzeitig zu vermeiden.

5 Konzeption eines Berechtigungskonzeptes basierend auf Dokumentenklassifikation

5.1 Umsysteme

Um das Konzept für die Datenklassifizierung besser nachvollziehen zu können, wird in diesem Kapitel das geplante Gesamt-Berechtigungskonzept erläutert. Daten-Labeling und Klassifikation werden ergänzend zu diesem Konzept eingesetzt.

In der Stadtverwaltung fiel mit der flächendeckenden Einführung von SharePoint gegen 2025 die grundlegende ACL-Berechtigungsstruktur weg. Neue Mitarbeitende erhalten keine direkte Zuteilung von Berechtigungen mehr, sondern eine oder mehrere Rollen. Diesen Rollen sind wiederum bestimmte Zugriffsrechte zugeteilt. Rollen werden anhand des Jobs der Mitarbeitenden vergeben. Das bedeutet, dass die Position «Support-Mitarbeitender» unabhängig von der Person, die diese Stelle besetzt, immer dieselben Rollen zugewiesen hat. Somit kann anhand der im Arbeitsvertrag festgelegten Aufgaben jederzeit abgeleitet werden, welche Rolle den einzelnen Mitarbeitenden zugewiesen ist.

Durch die im Kapitel 3.2 beschriebene Verlagerung von Daten in die Microsoft Cloud bzw. SharePoint ist das Gesamtkonzept auf die Strukturen von SharePoint ausgerichtet. Bei diesen Strukturen handelt es sich noch um Entwürfe in der Anfangsphase. Es ist mit Änderungen zu rechnen.

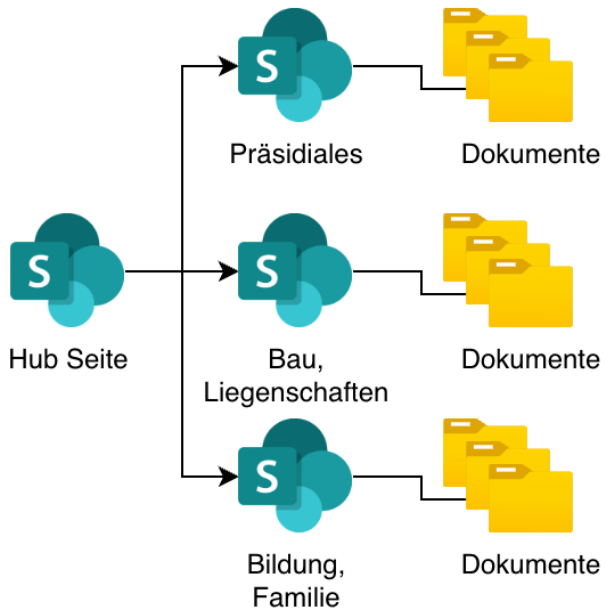

Abbildung 3: Aufbau SharePoint-Ablage (Entwurf)

In der Hierarchie zuoberst steht eine sogenannte Hub-Seite. Diese Hub-Seite kann über den Webbrowser von jedem beliebigen Endgerät aufgerufen werden. Sie dient als Einstiegspunkt, um auf die nächste Ebene der Datenhierarchie zu gelangen. Diese tiefere Ebene beinhaltet weitere SharePoint Seiten. Jedes Ressort der Stadtverwaltung besitzt eine eigene Seite. Über die Hub Seite kann, sofern die entsprechende Rolle vorhanden ist, auf die für das ressortspezifische Seite gelangt werden. Innerhalb der Ressort-Seite befinden sich erneut Seiten, diesmal aufgeteilt nach den Abteilungen, die in dem jeweiligen Ressort vorhanden sind. Zuletzt gibt es innerhalb dieser Abteilungsseiten eine Dokumentenbibliothek, welche die Ordner und Dateien der jeweiligen Abteilungen enthalten.

Der Zugriff auf die einzelnen Seiten und Dokumentenbibliotheken wird jeweils über die bereits erwähnten Rollen geregelt.

Auf der Ebene der Dokumentenbibliotheken wird nun die Dateiklassifikation eingesetzt. Werden Dateien innerhalb von SharePoint abgespeichert, müssen diese mit einem Label versehen werden und klassifiziert sein.

5.2 Architektur und Komponenten

In der Stadtverwaltung werden zurzeit alle Endpunkte innerhalb der eigenen Netzwerkumgebung betrieben. Jeglicher Datenverkehr in das Internet und somit auch die Microsoft Cloud wird von der eigenen Firewall kontrolliert. Diese steuert und überwacht, welche Endgeräte auf welche Webseiten zugreifen und ob Daten hoch- bzw. heruntergeladen werden dürfen. Zusätzlich kann im internen Netzwerk auf lokale Ressourcen wie lokal gehostete Applikationen und Daten zugegriffen werden. Diese Verbindung von Endgeräten in das interne Netzwerk wird über einen VPN-Tunnel oder in den Gebäuden der Stadtverwaltung durch lokale WLAN-Netzwerke aufrechterhalten.

Zukünftig werden Endgeräte von Mitarbeitenden nicht mehr im internen Netzwerk angesiedelt sein. Die Möglichkeit, einen VPN-Tunnel zum internen Netzwerk aufzubauen, wird abgeschafft und die lokale WLAN-Verbindung wird nicht mehr als internes Netzwerk behandelt. Ressourcen, welche sich im lokalen Rechenzentrum befinden, werden entweder komplett in die Microsoft Cloud verlagert oder über Reverse Proxy-Anbindung bereitgestellt. Auf den Endgeräten selbst wird die umfangreiche Endpoint-Sicherheitslösung von Microsoft, Microsoft Defender for Endpoint, bereitgestellt. Die Endgeräte werden über die Mobile Device Management-Lösung Micro-

soft Intune verwaltet. Diese funktioniert unabhängig vom Netzwerk des Endgerätes, solange dieses mit dem Internet verbunden ist. So können Applikationen aber auch Konfigurationen des Endgerätes verwaltet werden, ohne dass eine interne Verbindung nötig ist.

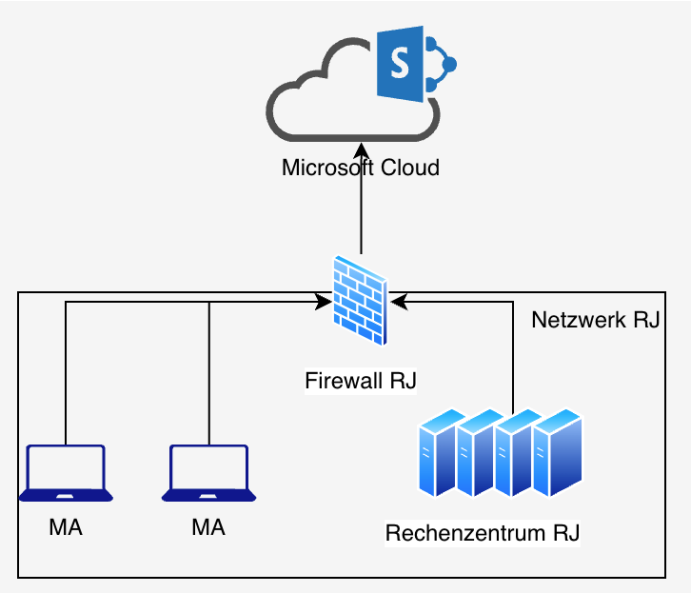

Abbildung 4: Ist-Architektur

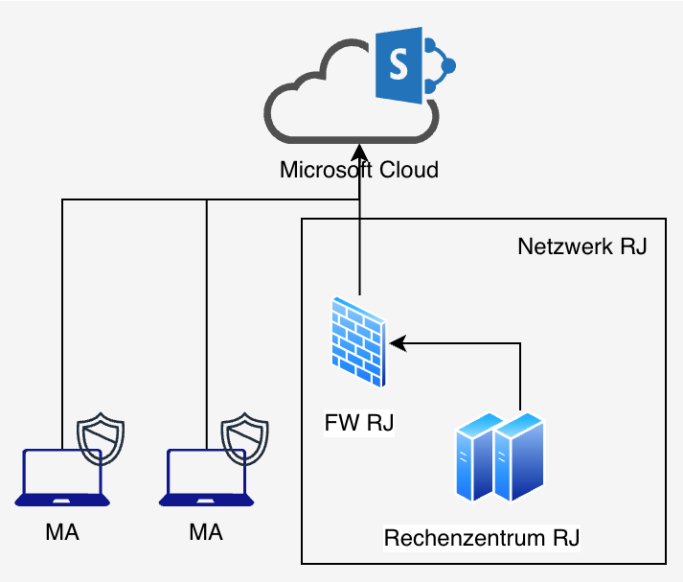

Abbildung 5: Soll-Architektur

Die Datenklassifikationslösung siedelt sich im Idealfall auf den Endgeräten selbst und innerhalb der Microsoft Cloud an und stärkt somit die Datensicherheit dort, wo am meisten mit diesen gearbeitet wird. Die Soll-Architektur wird bereits auf einigen Endgeräten erfolgreich getestet und ist in den Schulen der Stadt Rapperswil-Jona bereits teilweise im Einsatz.

5.3 Ziele

Im Folgenden sind Ziele definiert, die mit dem Konzept zur Datenklassifikation erreicht werden sollen.

Verhinderung des ungewollten Abflusses von Daten

Einer der wichtigsten Aspekte des neuen Klassifikationskonzeptes ist, den ungewollten Abfluss von Informationen zu verhindern. So muss eingeschränkt werden können, wie ein Dokument mit einer Klassifikation von Mitarbeitenden und Dritten verwendet werden darf.

Für die Klassen sollen folgende Einschränkungen für das Teilen von Daten gelten:

Geheim

Daten dürfen nur mit spezifizierten internen Personen geteilt werden, welche nachgewiesenen Bedarf für den Zugriff haben. Externes Teilen ist in der Regel verboten. Ausnahmen müssen durch ein Genehmigungsverfahren erlaubt werden. Der Versand dieser Daten per Mail, Messenger oder der Upload auf Webseiten wird blockiert. Dabei müssen die Dateien selbst und deren Inhalt geschützt werden. So wird auch das Kopieren von Inhalten und Einfügen in eine andere, niedriger eingestufte Datei oder Software verhindert. Zusätzlich werden Bildschirmaufnahmen und das Teilen des Bildschirms durch entsprechende Software (z.B. Screenshot, Microsoft Teams) blockiert. Dem Fotografieren der Datei mit Fotoapparaten und Handys wird mit einem Wasserzeichen entgegengewirkt. Das Wasserzeichen zeigt die Klassifizierung und den Namen der Person, welche die Datei geöffnet hat, an.

Vertraulich

Internen Zugriff erhalten nur Personen, welche entsprechenden Bedarf für die Verarbeitung dieser Daten haben. Dateien können intern nur mit Personen mit entsprechender Freigabestufe und Arbeitsbedarf geteilt werden. Externes Teilen ist nur mit vertrauenswürdigen Partnern und nur, wenn es für die Geschäftsbeziehungen notwendig ist, erlaubt. Der Versand der Datei per Mail an nicht freigegebene E-Mail-Adressen wird blockiert. Das Kopieren des Inhaltes wird ebenfalls blockiert. Zusätzlich wird bei der geöffneten Datei ein Wasserzeichen eingeblendet. Das Wasserzeichen zeigt die Klassifizierung und den Namen der Person, welche die Datei geöffnet hat, an.

Intern

Interne Dateien sind für alle internen Personen freigegeben. Dateien dürfen in bestimmten Fällen mit externen Personen geteilt werden. Der Massenversand an viele Empfänger wird blockiert. Zusätzlich wird bei der geöffneten Datei ein Wasserzeichen eingeblendet. Das Wasserzeichen zeigt die Klassifizierung und den Namen der Person, welche die Datei geöffnet hat, an.

Öffentlich

Als öffentlich eingestufte Daten sind für alle internen und externen Personen frei zugänglich. Daten können ohne Einschränkung geteilt werden.

Protokollierung und Überwachung

Jegliche Interaktionen mit Daten werden aufgezeichnet, protokolliert und überwacht. Dies geschieht unabhängig von der Klassifikation und in Echtzeit. Folgende Punkte müssen mindestens protokolliert werden:

- Basisinformationen über den Zugriff (Wer, wann, welche Datei, welche Interaktion?)
- Upload auf Webseiten oder Clouddiensten.
- Versuch, Informationen in einem nicht autorisierten Browser einzufügen.
- Kopieren oder Verschieben der Datei auf einen USB-Stick oder auf alternative externe Speichermedien
- Kopieren oder Verschieben auf einen Netzwerk-Speicher (z.B. privates NAS)
- Ausdrucken der Datei oder des Inhaltes
- Kopieren oder Verschieben der Datei oder deren Inhalts über Remote Desktop-Technologien
- Kopieren der Datei oder deren Inhalts in die Zwischenablage
- Zugriff durch Applikationen
- Versenden der Datei oder deren Inhalts per Mail oder alternative Nachrichtentools

Die Protokolle werden überwacht, um Auffälligkeiten oder Verstösse frühzeitig zu erkennen. Der Verstoss gegen Einschränkungen oder der Versuch, solche zu umgehen muss einen Alarm auslösen, welcher durch den Datenschutzbeauftragen bearbeitet werden muss.

Einschränkungen von Interaktionen mit Daten

Aufgrund der Klassifikation und der festgelegten Regeln sollen für Dateien und deren Inhalt Einschränkungen gelten. Die folgende Tabelle zeigt auf, welche Einschränkungen auf welche Klassifikation zutreffen. Für mit «X» markierte Felder gilt die in der Spalte «Einschränkung» definierte Einschränkung. Einzelne Felder enthalten zusätzlich weitere Informationen zu diesen Einschränkungen. Trifft eine Einschränkung nicht auf eine Klassifikation zu, wird dies mit «Keine Einschränkung» gekennzeichnet.

Einschränkung	**Geheim**	**Vertraulich**	**Intern**	**Öffentlich**
Zugriff durch Personen (Unabhängig von Klasse immer über Zugriffsrechte geregelt)	X Nur mit Freigabestufe. MFA notwendig	X MFA notwendig	X	Keine Einschränkung
Öffnen der Datei nur in vorgegebener Software	X	X	Keine Einschränkung	Keine Einschränkung
Zugriff nur aus bestimmten Ländern (Geoblocking)	X Nur aus der Schweiz	X Nur aus der Schweiz	X Zugriff nur aus der EU und der Schweiz	X Zugriff nur aus der EU und der Schweiz. Falls auf Webseite publiziert, weltweit.
Einschränkung der zugreifenden Endgeräte	X Nur verwaltete Endgeräte mit hoher Sicherheitsstufe	X Nur verwaltete Endgeräte	X Verwaltete Endgeräte. Zugriff auf privaten Endgeräten nur über verwaltete Plattform.	Keine Einschränkung

Verschlüsselung der Datei	X at Rest, in Motion	X at Rest, in Motion	X at Rest, in Motion	Keine Einschränkung
Upload auf Webseiten und Cloudumgebungen	X	X	X Nur wenn freigegeben.	Keine Einschränkung
Teilen von Daten (Siehe Abschnitt **Ungewollten Abfluss von Daten verhindern)**	X	X Nach Freigabe, per Mail an autorisierte Personen Freigabe muss terminiert sein.	X Nur per Mail	Keine Einschränkung
Kopieren auf externe Speichermedien (USB, Festplatten)	X	X	X Kann übersteuert werden.	Keine Einschränkung
Datei ausdrucken	X	X	Keine Einschränkung	Keine Einschränkung
Kopieren von Inhalt in die Zwischenablage	X	X	Keine Einschränkung	Keine Einschränkung

Tabelle 2: Einschränkungsmatrix

Verschlüsselung von Informationen

Dateien sollen aufgrund ihrer Klassifikation verschlüsselt werden können. So soll verhindert werden, dass Informationen auf unautorisierten Geräten geöffnet und bearbeitet werden. Das Dokument darf nur von autorisierten Usern geöffnet werden können. Die Verschlüsselung muss unabhängig vom Speicherort sein, die Datei selbst ist verschlüsselt. Eine allfällige Verschlüsselung des Speichermediums ist zusätzlich möglich, aber nicht zwingend. Die Verschlüsselung muss

«at Rest» (z.B. in der Cloudablage gespeichert) und «in Motion» (z.B. beim Versenden per Mail) verschlüsselt sein.

Applikationsunabhängige Klassifikation

Applikationen, welche in der Stadtverwaltung Rapperswil-Jona eingesetzt werden, um Dateien zu bearbeiten, müssen die Klassifikationen unterstützen. So soll beispielsweise eine Worddatei, welche als «Geheim» klassifiziert ist, auch nach dem Speichern als PDF-Datei und dem Öffnen in einem entsprechenden Editor als «Geheim» klassifiziert sein.

Speicherortunabhängige Klassifikation

Klassifikationen müssen ortsunabhängig immer beibehalten werden. Wird ein Dokument in einem lokalen Verzeichnis mit einem Label und somit einer Klassifikation belegt, muss diese Klassifikation auch nach dem Verschieben in einen anderen lokalen Speicher oder einen Cloud-Speicher bestehen bleiben. Nur so kann sichergestellt werden, dass die Klassifizierung nicht umgangen werden kann.

Wird eine Datei kopiert, so muss auch die Kopie dieser Datei dieselbe Klassifikation aufweisen. Auch das Umbenennen der Datei darf keinen Einfluss auf die Klassifikation haben.

6 Implementierungsstrategie und Proof of Concept

6.1 Implementierungsstrategie

Um Datenklassifikation und damit verbundene Sicherheitsvorschriften effektiv in der Stadtverwaltung umzusetzen, wird in diesem Kapitel eine Implementierungsstrategie aufgezeigt. Vorgehensweisen bei der Implementierung gibt es so viele, wie es Datenklassifikations- und DLP-Systeme gibt. Daher ist die hier aufgezeigte Strategie allgemein gehalten. Ein spezifisches Vorgehen zur Implementierung von Microsoft Purview ist im Kapitel 6.5 zu finden.

6.2 Vorbereitungsphase

6.2.1 Wieso wird eine Implementierungsstrategie benötigt?

Datenklassifikations- und DLP-Systeme sind äusserst komplex und bestehen meist aus vielen verschiedenen Diensten, welche ineinandergreifen. Es sind jedoch nicht nur die technischen Aspekte, die eine Implementierung zu einem vielschichtigen Prozess machen. Es ist wichtig, dass diese Lösungen auch auf einer strategischen Unternehmensebene eingeführt werden (Rogowski, 2013). Dies hilft, dass sich auch das gesamte Unternehmen an den neuen Vorschriften hält und diese auch effektiv umgesetzt werden können. Für die Stadtverwaltung ist bereits klar, dass mit der Microsoft-Lösung unternehmensweit DLP eingeführt wird. Die Unterstützung auf Management- und Strategieebene ist somit bereits gesichert und es muss keine Lösung mehr evaluiert werden.

Es sind jedoch noch weitere Schritte für eine erfolgreiche Implementierung nötig. Für das Vorgehen orientiert sich diese Strategie an der Empfehlung von Microsoft und deren Guide «Deploy an Information protection Solution with Microsoft Purview» (Bailey et al., 2023).

6.2.2 Identifikation der Stakeholder

Mitarbeitende	Die Mitarbeitenden der Stadtverwaltung sind am meisten von den Änderungen betroffen, welche die Datenklassifikation mit sich bringt. Die Arbeit mit einem Grossteil der Daten der Stadtverwaltung wird täglich von ihnen durchgeführt.

Informatikdienst	Der Informatikdienst der Stadtverwaltung ist zuständig für die technische Umsetzung der Klassifikationslösung. Zusätzlich ist der Informatikdienst über den gesamten Zyklus der Lösung für die Überwachung, Aktualisierung und stetige Verbesserung zuständig.
Stadtrat	Der Stadtrat mit dem Stadtpräsidenten leitet die Stadtverwaltung. Sie stellen die Ressourcen für die Beschaffung, Umsetzung und den Unterhalt bereit. Sie sogen für die Datensicherheit und Effizienz der Stadtverwaltung.
Ressortleitung	Die Ressortleitungen sind für die einzelnen Ressorts der Stadtverwaltung verantwortlich und leiten diese.
Bürgerinnen und Bürger	Die Bürgerinnen und Bürger der Stadt Rapperswil-Jona nutzen städtische Dienste und fordern eine datenschutzgerechte Verwaltung ihrer anvertrauten Daten.
Datenschutzbeauftragter	Der Datenschutzbeauftragte überwacht die Einhaltung des schweizerischen Datenschutzgesetzes (DSG)

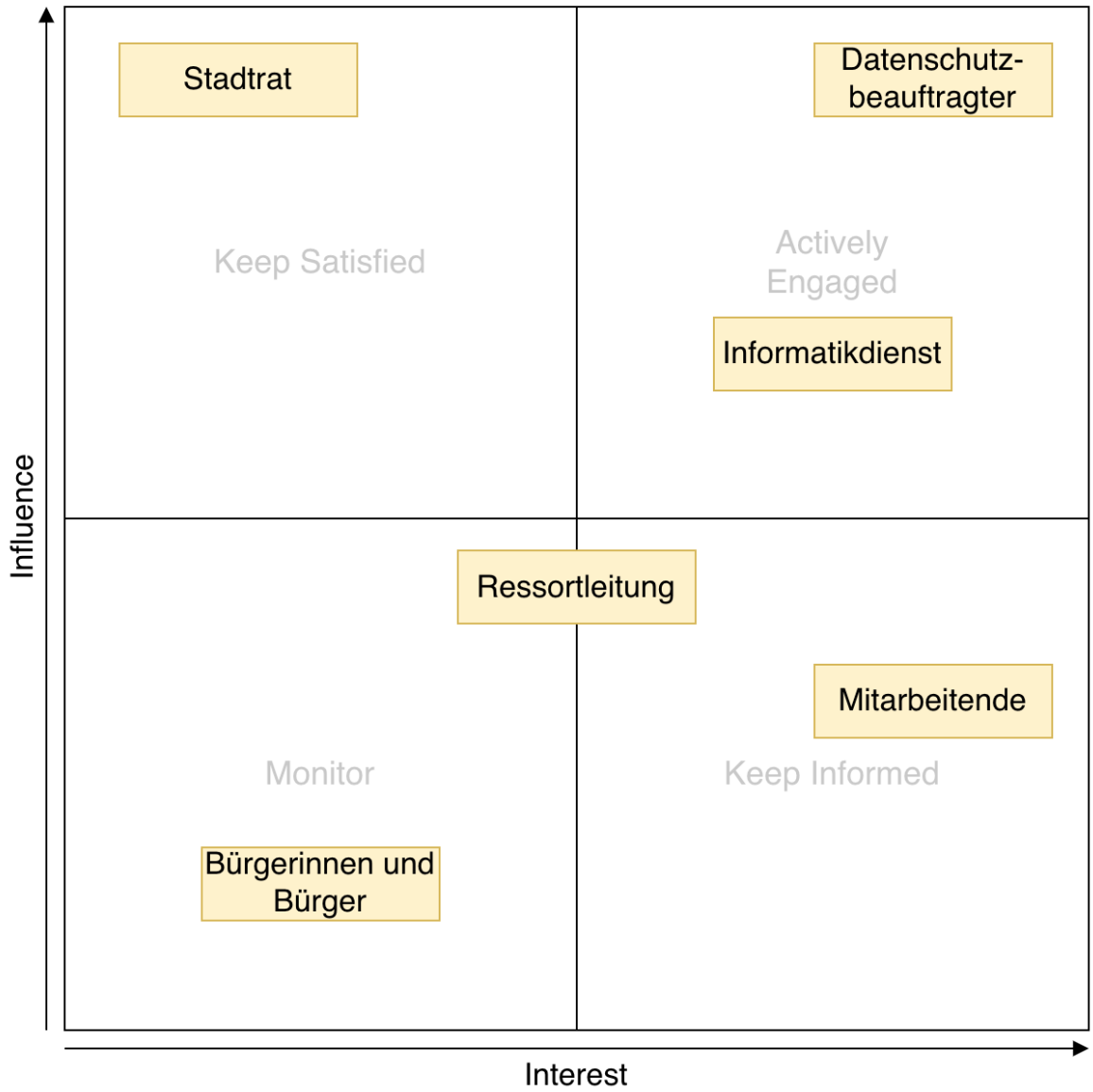

Abbildung 6: Stakeholder Mapping für Datenklassifikation und DLP

6.2.3 Identifikation der zu schützenden Daten

Wie bereits beschrieben, ist es wichtig, dass die Stadtverwaltung ihre eigenen Daten kennt. Es sollte jederzeit klar sein, welche sensitiven Daten sich wo im System befinden. Ohne diese essenziellen Informationen ist es nicht möglich, effektive Regeln zu erstellen und diese durchzusetzen. Die gängigen DLP-Systeme bieten Funktionalitäten, um bestehende Datenspeicher nach spezifischen Informationen zu durchsuchen und Berichte zu generieren. Dieser Vorgang kann bereits früh gestartet werden, da dabei nur Daten nach sensitiven Informationen durchsucht werden und keine Einteilung in Klassen stattfindet.

Wenn die ersten Informationen in den Dateiablagen entdeckt werden, kann mit der Anpassung und der Verbesserung der Erkennungsmechanismen begonnen werden. Je nach Mechanismus

(Exact Data Match, Machine Learning usw.) ist es nötig, diesen über eine Zeitspanne immer wieder anzupassen. So können für Machine Learning beispielsweise immer wieder neue Muster zum Trainieren verwendet werden.

6.3 Einführung

6.3.1 Klassifikation der Daten

Wenn eine Übersicht über die Daten in der Stadtverwaltung vorhanden ist, kann festgelegt werden, welche Informationen wie geschützt werden müssen. Damit dies nicht einzeln für jede Datei gemacht werden muss, sollten diese klassifiziert werden. Im Fall der Stadtverwaltung werden die Klassen **geheim**, **vertraulich**, **intern** und **öffentlich** verwendet. Diese Klassen richten sich nach dem Leitfaden des Kantons St. Gallen wie in Kapitel 4.1 beschrieben. Die Einteilung muss einheitlich, klar definiert und dokumentiert sein. Dies ist wichtig, da jederzeit nachvollziehbar sein muss, wieso eine Information in eine der Klassen eingeteilt wurde.

Wenn die Klassen definiert sind und Informationen möglichst eindeutig in diese Klassen eingeteilt werden können, wird die Klassifikation im nächsten Schritt angewendet. Für die Stadtverwaltung wird eine Mischung aus automatischer und manueller Klassifikation angewendet. Aufgrund des Voranschreitens der Implementierung der Klassifikation verändern sich der Anteil von automatischer und manueller Klassifikation sowie deren Auswirkung auf die Dateien.

Während der Einführungsphase werden Dokumente nicht automatisch klassifiziert und die Klassen verursachen keine Einschränkungen bei Interaktionen mit den Daten. Die Klassifikation erfolgt in dieser Phase ausschliesslich manuell. So wird sichergestellt, dass erste ausgewählte Mitarbeitende das Klassifikationssystem kennenlernen und testen können. Diese Phase dauert etwa sechs Monate, wobei mit voranschreitender Zeit zunehmend Mitarbeitende in das System eingeführt werden. In dieser Zeit werden in regelmässigen Abständen die klassifizierten Daten überprüft und Rückmeldung der Mitarbeitenden eingeholt bzw. gegeben. Sollten grössere Unklarheiten oder Probleme auftreten, kann sich diese Phase verlängern.

Nach der Einführungsphase wird abteilungsweise die Klassifizierung eingeführt. Ab dann wird jeder neuen oder bearbeiteten Datei immer standardmässig die Klasse **intern** zugewiesen. Die Mitarbeitenden haben jedoch immer die Möglichkeit, diese standardmässige Klassifikation anzupassen. Einem neuen Dokument muss jedoch immer eine Klasse zugewiesen sein. Zusätzlich werden erste automatische Mechanismen eingeschaltet. Diese klassifizieren jedoch nicht selbst, sondern bieten den Mitarbeitenden lediglich Vorschläge für eine mögliche Klassifizierung

aufgrund des erkannten Inhaltes an. Diese Erkennung muss nicht von Beginn an alle Informationen abdecken. Vielmehr ist es wichtig, diese über die Zeitspanne der schrittweisen Einführung immer weiter auszubauen und zu verbessern. Es wird weiterhin keine Einschränkungen basierend auf den Klassen geben. Ziel ist es, alle Mitarbeitenden mit dem Klassensystem vertraut zu machen und ihnen erste Möglichkeiten zu geben, Daten selbst zu klassifizieren oder die bestehende Klassifizierung **intern** aktiv zu hinterfragen. Auch diese Phase dauert nochmals sechs Monate und kann, falls nötig, verlängert werden.

Nach diesen beiden Phasen sollten alle Mitarbeitenden bereits erste Erfahrungen mit der Klassenzuteilung besitzen. DLP-Software bietet die Möglichkeit, zu überwachen, wie und von wem Klassen angewendet wurden. Dies hilft dabei, Mitarbeitende, welche bis zu diesem Zeitpunkt nicht oder nur wenig klassifiziert haben, zu erkennen und aktiv auf diese zuzugehen. Jede Änderung der Standardklasse oder der vorgeschlagenen Klasse dient gleichzeitig auch als Rückmeldung für die automatischen Mechanismen zur Inhaltserkennung. Durch diese Rückmeldungen der Mitarbeitenden können Fehler in der Konfiguration oder in den Mechanismen erkannt werden und entsprechende Änderungen vorgenommen werden.

6.3.2 Einschränkungen durch Klassen

Wenn die Mitarbeitenden mit dem Verwenden der Klassen vertraut sind, werden die Einschränkungen, welche in Kapitel 5.3 definiert wurden, eingeschaltet. Je nach Schweregrad können alle auf einmal oder jeweils einzelne Einschränkungen eingeführt werden. Dadurch können Rückmeldungen von Mitarbeitenden besser abgefangen und mögliche Komplikationen früh erkannt werden. Durch die schrittweise Einführung der Einschränkungen können sich die Mitarbeitenden einfacher an diese gewöhnen. Ziel ist es, dass die definierten Einschränkungen alle eingeschaltet sind. Anpassungen bleiben jedoch vorbehalten, sollte es zu starken Einschränkungen bei der täglichen Arbeit kommen.

6.3.3 Schulung der Betroffenen

Ein wichtiger, wenn nicht der wichtigste, Bestandteil der Einführung der Klassifikation ist die Schulung der von der Klassifikation betroffenen Mitarbeitenden. Ziele der Schulungen sind:

- Vermittlung von Verständnis für die verschiedenen Klassifikationsstufen (öffentlich, intern, vertraulich, geheim) und deren Auswirkung auf Daten
- Bewusstsein für die Bedeutung der Datenklassifikation und der Datensicherheit.
- Korrekte und richtliniengemässe Anwendung der Klassen.

Die Schulungen umfassen die Themen:

- Grundlagen des Datenschutzes und die Notwendigkeit der Datenklassifikation.
- Detaillierte Erläuterung der einzelnen Klassifikationsstufen und wann diese zum Einsatz kommen.
- Praktische Anleitung zur korrekten Klassifizierung von Dokumenten.
- Richtlinien für die Handhabung und Weitergabe klassifizierter Dokumente.
- Verfahrensweisen bei der Entdeckung und Korrektur von Fehlklassifikationen, beispielsweise das Angeben von Gründen.

Um den Mitarbeitenden das Wissen möglichst effizient zu vermitteln, werden zu Beginn einige Präsenzschulungen mit ausgewählten Mitarbeitenden durchgeführt. So wird der direkte Austausch gefördert und ein initialer Wissenstransfer kann stattfinden. Zudem werden auf der verwaltungseigenen Wissensplattform (MySupport) Onlineressourcen zur Verfügung gestellt, auf welche jederzeit zugegriffen werden kann. Diese Ressourcen beinhalten dieselben Informationen, welche in den Präsenzschulungen vermittelt werden. Der Fokus liegt klar auf diesen Onlineressourcen. Sie dienen als Informations- und Dokumentations-Hub für alle Mitarbeitenden. Der Inhalt wird laufend ergänzt und aktualisiert.

Essenzielle Ressourcen sind vor der Einführung zu erstellen, damit diese zusammen mit dem Ausrollen der ersten Funktionalitäten bereit sind. Während eines Kick-off-Termins können die ersten Mitarbeitenden informiert und über die wichtigsten Funktionen geschult werden. Wichtig ist, dass die Mitarbeitenden während und nach der Einführungsphase bei Unklarheiten jederzeit auf die wichtigsten Informationen zurückgreifen können.

6.4 Nach der Einführung

Nach der Einführung der Dokumentenklassifikation ist es wichtig, die Effektivität und Funktionalität der getroffenen Massnahmen sicherzustellen. Dafür muss eine stetige Überwachung, Bewertung und Anpassung der Klassifikationen sichergestellt werden. Folgend werden die dafür zu treffenden Massnahmen erläutert.

6.4.1 Monitoring und Audit

Wie bereits erwähnt ist das Monitoring schon während der Einführung ein wichtiger Aspekt zur Überwachung und Verbesserung der Datenklassifikation durch die Mitarbeitenden. Auch nach der Einführung ist das Monitoring weiterhin wichtiger Bestandteil. Hinzu kommt das Auditing, um die Effektivität der Klassifikation zu gewährleisten. Es muss überprüft werden, ob bestehende

Klassifikationsrichtlinien durch die Mitarbeitenden eingehalten werden und den organisatorischen Anforderungen entsprechen. Durch Audits muss in regelmässigen Abständen überprüft werden, ob die Genauigkeit der automatischen Klassifikation den Anforderungen genügt und dass es zu möglichst wenigen Fehlklassifikationen kommt.

6.4.2 Anpassung und Optimierung

Mit den gewonnenen Erkenntnissen aus dem Monitoring und Auditing kann das System angepasst und optimiert werden. Die Klassifikationsrichtlinien können aufgrund der sich verändernden organisatorischen Bedürfnisse weiterentwickelt werden. Es können weitere Mechanismen zur automatischen Klassifikation eingeführt und bestehende überarbeitet werden. Mit einer wachsenden Anzahl an klassifizierten Dateien und klassifizierenden Mitarbeitenden können allfällige Machine Learning-Algorithmen besser trainiert und optimiert werden. So kann deren Effizienz gesteigert und die Fehleranfälligkeit verringert werden.

Ein weiter wichtiger Punkt ist der Unterhalt und die Weiterentwicklung des eingesetzten Systems. Die Software muss gepflegt und auf dem neusten oder möglichst aktuellen Stand gehalten werden. Erscheinen neue Features oder bestehende Features und Funktionen ändern sich, so ist festzustellen, wie sich dies auf das System der Stadtverwaltung auswirkt.

6.4.3 Schulung und Sensibilisierung

Nebst der Schulung während der Einführung ist auch eine kontinuierliche Wissensvermittlung nach der Einführungsphase wichtig. Mitarbeitende müssen weiterhin für die Datensicherheit sensibilisiert werden. Damit wird die Bedeutung der Datenklassifikation aufrechterhalten und die Mitarbeitenden halten sich eher an die Richtlinien, da sie verstehen, wieso diese vorhanden sind.

Zusätzlich ist es wichtig, regelmässiges Feedback der Mitarbeitenden aber auch der anderen Stakeholder einzuholen. Durch dessen Auswertung kann die Effektivität und die Benutzerzufriedenheit evaluiert werden.

6.5 Pilot /Proof of Concept

In den vorhergehenden Kapiteln wurden verschiedene organisatorische und technische Aspekte der Datenklassifikation erläutert. Es wurde bewusst eine technologieunabhängige Sichtweise verwendet, um eine konzeptionelle Grundlage für die Einführung verschiedenster Lösungen abzudecken. In diesem Kapitel werden nun anhand einer konkreten Softwarelösung diese konzeptionellen Schritte praktisch umgesetzt, um einen Proof of Concept für die Stadtverwaltung Rapperswil-Jona zu erstellen.

6.5.1 Microsoft Purview

Der Proof of Concept wird mit anhand der Microsoft Lösung Microsoft Purview umgesetzt. Bei Microsoft Purview handelt es sich um eine Suite von Softwarelösungen, um Organisationsdaten zu schützen, zu verwalten und zu steuern (Mazzoli et al., 2024). Purview soll Unternehmen dabei unterstützen, ihre Daten zu kennen, zu managen und zu schützen, sensitive Daten zu klassifizieren und die Compliance zu verbessern.

Microsoft Purview wird für den Proof of Concept verwendet, weil die Stadtverwaltung bereits zu grossen Teilen auf die Microsoft 365 Softwaresuite setzt. Unter anderem gehören dazu die Lizenzmodelle Microsoft 365 A3, E3, A5 und E5. Insgesamt sind über 5000 Benutzer mit diesen Lizenzen ausgestattet. Diese Lizenzen beinhalten alle die vollständig oder zumindest Teile von Microsoft Purview, welche für die Dokumentenklassifikation benötigt werden. Somit könnte die Dokumentenklassifikation ohne zusätzliche Lizenzkosten realisiert werden.

Ein weiterer grosser Vorteil der Microsoft-Lösung ist die ausgezeichnete Integration mit den weiteren Produkten der Microsoft 365 Suite. Es werden keine zusätzliche Software, Erweiterungen oder Plugins für die Klassifikation benötigt. Auch bieten einige Softwarehersteller eine Integration an. So können beispielsweise in den Adobe-Produkten PDF-Dokumente direkt per Dropdown-Menü klassifiziert werden.

All diese Punkte machen Microsoft Purview zu einer vielversprechenden Lösung im Bereich der Datenklassifikation für die Stadtverwaltung. In diesem Proof of Concept wird nun überprüft, ob sich dies auch praktisch umsetzen lässt. Aufgrund des Umfangs wird sich dieser Proof of Concept auf die wichtigsten Funktionen beschränken.

6.5.2 Setup

Das Setup wurde anhand des Setup Giude von Microsoft durchgeführt (Microsoft Purview Information Protection Setup Guide | MIP Setup Guide, o. J.).

Voraussetzung für Microsoft Purview ist ein aktiver Microsoft 365 Tenant und eine der bereits erwähnten Lizenzen von Microsoft 365. Zusätzlich wird für das Setup eine globale Administratorenrolle benötigt. Tenant sowie Rollen sind bereits vorhanden, da diese auch für andere Produkte, welche bereits im Einsatz sind, verwendet werden. Daher kann sofort mit dem Einrichten von Labels begonnen werden.

6.5.3 Labeling und Klassifizierung

In Purview wird für Klassen der Begriff Sensitivity Label verwendet (Bailey, Robertson, et al., 2024). Es wird demnach für jede der definierten Klassen ein solches Label erstellt. Dies kann über das Purview-Portal durchgeführt werden. Beim Erstellen eines Labels werden ein Name (öffentlich, intern, vertraulich, geheim) und eine Beschreibung, wofür dieses Label steht, angegeben. Es kann zudem festgelegt werden, auf welche Elemente diese Labels angewendet werden können. Möglichkeiten, die Purview bietet, sind:

- Files (Office Files, PDFs, usw.)
- E-Mails (Nachrichten in Microsoft Outlook)
- Meetings (Kalendereinträge und Meetings in Outlook und Teams)
- Gruppen und Seiten (Microsoft Teams-Gruppen und SharePoint-Seiten)
- Schematized Data Assets (SQL, Azure SQL, AWS RDS, usw.)

Für diesen PoC sollen die wichtigsten Elemente Files, E-Mails und Meetings klassifiziert werden. Gruppen und Seiten werden aufgrund des Umgangs weggelassen und Schematized Data Assets befinden sich noch in einer Preview-Phase und werden nicht eingesetzt, bis diese vorüber ist.

Anschliessend wird festgelegt, wie sich diese Sensitivity Labels auf das Element auswirken, welchem diese zugewiesen werden (Bailey, Koenen, Cross, et al., 2024). Es können zwei Sicherheitseinstellungen vorgenommen werden:

- Zugriffsberechtigung (Verschlüsselung)
- Content Markings (Headers, Footers, Wasserzeichen)

Für den PoC wird bei allen Klassen ein Wasserzeichen angewendet, um vereinfacht darzustellen, ob die Klassifikation funktioniert hat. Jede Klasse soll ein Wasserzeichen erhalten, bei dem die

Klasse und der Benutzername der Person dargestellt wird, welche das Dokument bearbeitet. Dafür kann die eine Variable eingebettet werden, welche auf Benutzerattribute verweist.

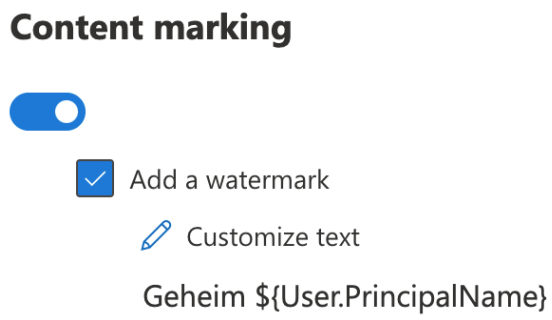

Abbildung 7: Content Marking mit Wasserzeichen

Mit diesen Einstellungen entsteht ein Wasserzeichen wie in Abbildung 8 gezeigt. Für die anderen Klassen können diese Wasserzeichen entsprechend angepasst oder auch entfernt werden. Wasserzeichen sind für geheime und vertrauliche Dokumente vorgesehen.

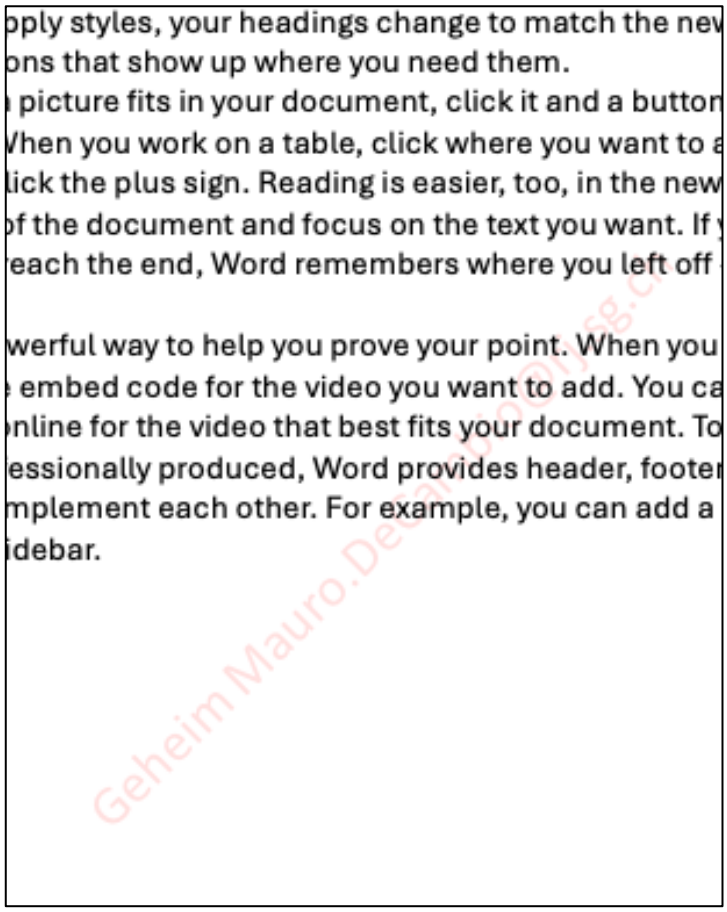

Abbildung 8: Content Marking-Wasserzeichen in MS Word

6.5.4 Verschlüsselung und Zugriffsteuerung

Der zweite Aspekt ist die Zugriffssteuerung bzw. Verschlüsselung (Bailey, Koenen, Cross, et al., 2024). Hier wird festgelegt, wer das Dokument entschlüsseln darf und welche Aktionen damit ausgeführt werden dürfen. Je nach Anforderung können hier unterschiedliche Einstellungen gesetzt werden. Hier werden nur die für diesen PoC wichtigen Einstellungen behandelt. Für den PoC werden für die Klasse geheim folgende Berechtigungen gesetzt.

Berechtigung	**Testuser1** (Erstellt und arbeitet mit geheimen Dokumenten)	**Testuser2** (Soll die geheimen Dokumente nur lesen können)
View Content	X	X
View Rights	X	X
Edit content	X	
Save	X	
Print	X	
Copy and extract content	X	
Reply (Mailobjekt)	X	
Reply all (Mailobjekt)	X	
Forward (Mailobjekt)		
Edit rights		
Export Content		
Allow macros	X	X
Full Control		

Für alle anderen Mitarbeitenden und externen Personen bleiben Dateien mit dieser Klasse somit verschlüsselt. Die Berechtigungsmatrix bildet die vorgesehenen Einschränkungen so genau wie möglich ab. Weitere Einschränkungen für die Klassen werden zu einem späteren Zeitpunkt noch

eingerichtet. Anstelle von einzelnen Benutzern werden in der produktiven Umgebung Benutzergruppen zum Einsatz kommen.

Die Auswirkungen werden mithilfe von drei verschiedenen Benutzern getestet.

- Testuser1: Dieser Benutzeraccount erstellt das Dokument und klassifiziert dieses als geheim. Er kann die meisten Aktionen auf diesem Dokument durchführen.
- Testuser2: Dieser Benutzeraccount hat eingeschränkte Rechte, um das Dokument zu lesen, dieses jedoch weder zu bearbeiten noch den Inhalt zu extrahieren.
- Testuser3: Diesem Benutzeraccount werden keine Rechte zugeteilt, daher ist dieser auch nicht in der Matrix vorhanden. Er steht stellvertretend für alle Mitarbeitenden, welche keinen Zugriff auf das Dokument haben sollen.

Ein Word-Dokument wird vom Testuser1 erstellt und mit Testuser2 und 3 geteilt. Folgendes Verhalten wird bei Testuser2 beobachtet:

Das Dokument lässt sich öffnen und der Inhalt lesen. Das Dokument kann nicht bearbeitet, ausgedruckt und abgespeichert werden. Zusätzlich wird beim Versuch, einen Screenshot mit einer Bildschirmaufnahmesoftware (z.B. SnippingTool), der Bildschirm schwarz, sodass auf einer Aufnahme nichts zu erkennen ist. Gegen das Abfotografieren mit einem externen Gerät (z.B. mit dem Smartphone) ist das Wasserzeichen gross über die gesamte Seite zu sehen (vgl. Abbildung 8). Das Kopieren des Inhalts in die Zwischenablage wird unterbunden. Diese Einschränkungen bleiben auch beim Kopieren des Dokumentes erhalten. Das Dokument kann nicht ausgedruckt und exportiert werden. Die Berechtigungen wurden demnach korrekt angewendet.

Beim Testuser3, welcher keinerlei Berechtigungen auf mit geheim klassifizierten Dokumenten besitzt, lässt sich das Dokument nicht öffnen, da dieser nicht die erforderlichen Berechtigungen aufweist, um das Dokument zu entschlüsseln.

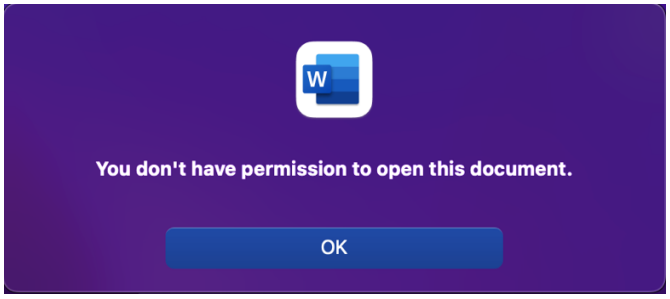

Abbildung 9: Meldung beim Versuch ein als geheim klassifiziertes Dokument zu öffnen.

Somit funktioniert die Steuerung des Zugriffs auf Dokumente unabhängig vom Speicherort.

6.5.5 Softwareintegration

Microsoft wird zusätzlich von einigen Drittanbietern, so zum Beispiel von Adobe Acrobat für PDFs unterstützt. (*Unterstützung für Microsoft Purview Information Protection in Acrobat*, 2023). Im folgenden Screenshot ist zu sehen, wie das Microsoft Purview-Label intern von Adobe Acrobat richtig erkannt und dargestellt wird.

Abbildung 10: Integration in Adobe Acrobat für PDF-Dateien

6.5.6 Automatische Klassifizierung und Vorschläge

Microsoft Purview bietet die Funktion, die Sensitivity Labels automatisch aufgrund des Inhalts eines Dokumentes entweder direkt anzuwenden oder einen Vorschlag für ein Label zu unterbreiten (Bailey, Koenen, Mazzoli, et al., 2024). Als Erkennungsmerkmal wird die schweizerische Sozialversicherungsnummer (AHV-Nummer) verwendet. Wenn zwei oder mehr AHV-Nummern innerhalb eines Dokumentes erkannt werden, soll das Label Vertraulich vorgeschlagen werden. Da die Stadtverwaltung höchstwahrscheinlich nur automatische Vorschläge und nicht die automatische Zuteilung von Klassen verwenden wird, wird in diesem PoC auch dieses Vorgehen genauer überprüft. Für die korrekte Erkennung des Inhaltes, besteht diesbezüglich jedoch kein Unterschied. Beim Vorschlag wird lediglich nochmals nach einer Benutzereingabe gefragt.

Damit Purview sensitive Informationen erkennt, gibt es die sogenannten Sensitive Information Types (SIT) (Fox & Koenen, 2024b). In diesen SIT können Informationen definiert werden, welche von der Stadtverwaltung als sensitiv eingestuft werden können. SIT bietet die musterbasierte Erkennung von Informationen. Sie eignen sich daher gut für die Erkennung von AHV-Nummern, da

diese immer demselben Muster folgen. Microsoft bietet eine Reihe an vordefinierten SITs, es können jedoch auch selbst Muster erstellt werden. Für die AHV-Nummer gibt es eine Vorlage, welche anhand folgenden Musters vorgeht (Fox et al., 2023):

- 13 Nummern
- Beginnt mit 3 Nummern 756
- Optionaler Punkt
- Vier Nummern
- Optionaler Punkt
- Vier Nummern
- Optionaler Punkt
- Zwei Nummern

Zusätzlich wird überprüft, ob innerhalb von 300 Zeichen vor oder nach der Nummer eines der folgenden Schlüsselwörter vorkommt:

- Ahv
- Ssn
- Pid
- Insurance number
- Personalidno#
- Social security number
- Personal id number
- Personal identification no.
- Insuranceno#
- Uniqueidno#
- Unique identification no.
- Avs number
- Personal identity no versicherungsnummer
- Identifikationsnummer
- Einzigartige identität
- Sozialversicherungsnummer
- Identification personnelle id
- Numéro de sécurité sociale

Der SIT erkennt eine höchstwahrscheinliche Übereinstimmung, wenn Muster und Schlüsselwort erkannt werden, und eine mittlere Übereinstimmung, wenn nur das Muster erkannt wird.

Das getestete Dokument wird als öffentlich klassifiziert. Erwartet wird, dass, sobald zwei AHV-Nummern in das Dokument eingefügt werden, die Klassifikation Vertraulich vorgeschlagen wird.

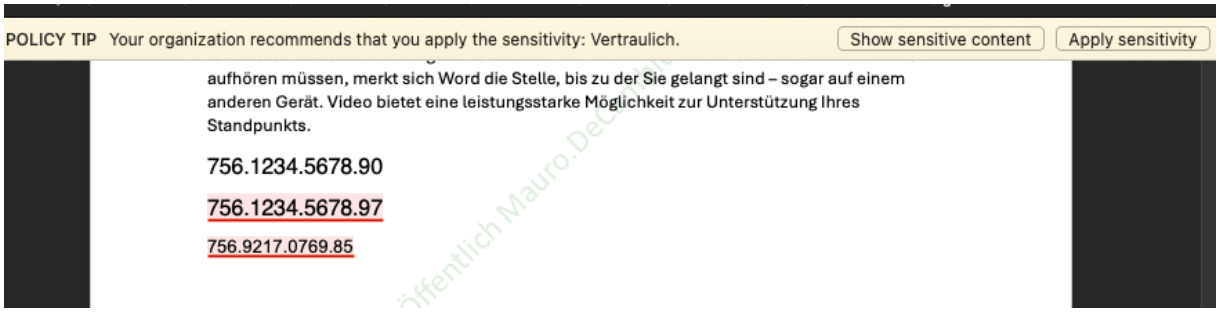

Abbildung 11: Vorgeschlagene Klasse aufgrund der AHV-Nummer

Wie in Abbildung 11 zu sehen ist, werden zwei der drei AHV-Nummern rot markiert und als Grund für den Vorschlag, die Klasse zu wechseln, angegeben. Da definiert ist, dass der Vorschlag erst ab zwei AHV-Nummern erscheinen soll, werden nur die zweite und die dritte AHV-Nummer markiert.

6.5.7 Weitere Einschränkungen mit DLP

Für weitere feinere Einschränkungen bei Dokumenten mit sensiblem Inhalt kann Microsoft Purview DLP verwendet werden. Wie der Name bereits vermuten lässt, lassen sich mit diesem Teil der Purview Suite mit einer Reihe von Regeln, welche direkt aufgrund des Inhaltes oder aufgrund der festgelegten Klasse angewendet werden, Einstellungen zu Data Leak Prevention vornehmen (Toelle, 2021). Da in der Stadtverwaltung geplant ist, dass alle Dateien einer Klasse zugewiesen werden müssen, werden DLP-Richtlinien nur aufgrund dieser Klassen angewendet. DLP wird für folgende Einschränkungen verwendet:

- Öffnen der Datei nur in Vorgegebener Software
- Teilen von Daten (Mail, Cloud usw.)
- Kopieren auf externe Speichermedien
- Upload auf Webseiten und Cloudumgebungen

Öffnen der Datei nur in vorgegebener Software

Für den PoC wird mit einer Richtlinie die Applikation notepad++.exe daran gehindert, als geheim oder vertraulich klassifizierte Dokumente zu öffnen. Auf einem Notebook wurde diese Richtlinie überprüft, indem das Dokument Dok1.docx als vertraulich klassifiziert wurde. Danach wurde versucht, das Dokument in der blockierten Applikation zu öffnen. Als Ergebnis erscheint die Meldung, dass das Öffnen des Dokuments in dieser Applikation nicht gestattet ist. Da diese Richtlinie nur für als vertraulich und geheim klassifizierte Dateien gilt, können öffentliche und interne Dateien weiterhin in der Applikation geöffnet werden.

Abbildung 12: Meldung beim Versuch, ein Dokument in einer blockierten App zu öffnen

Teilen von Daten (Mail, Cloud usw.)

Eine weitere wichtige Einschränkung, vor allem für die Klassen geheim und vertraulich, ist das Teilen mit externen Personen. Eine entsprechende DLP-Richtlinie kann eingerichtet werden, um zu kontrollieren, dass Dokumente mit einer der beiden Klassen weder per Mail verschickt noch direkt über die Teilen-Funktion in SharePoint oder OneDrive geteilt werden können. Wird eine E-Mail erstellt und eine Datei mit entsprechender Klassifikation als Anhang eingefügt, wird die E-Mail selbst in diese Klasse eingestuft. Zusätzlich wird eine entsprechende Meldung angezeigt und das Senden der Nachricht an externe Personen wird blockiert.

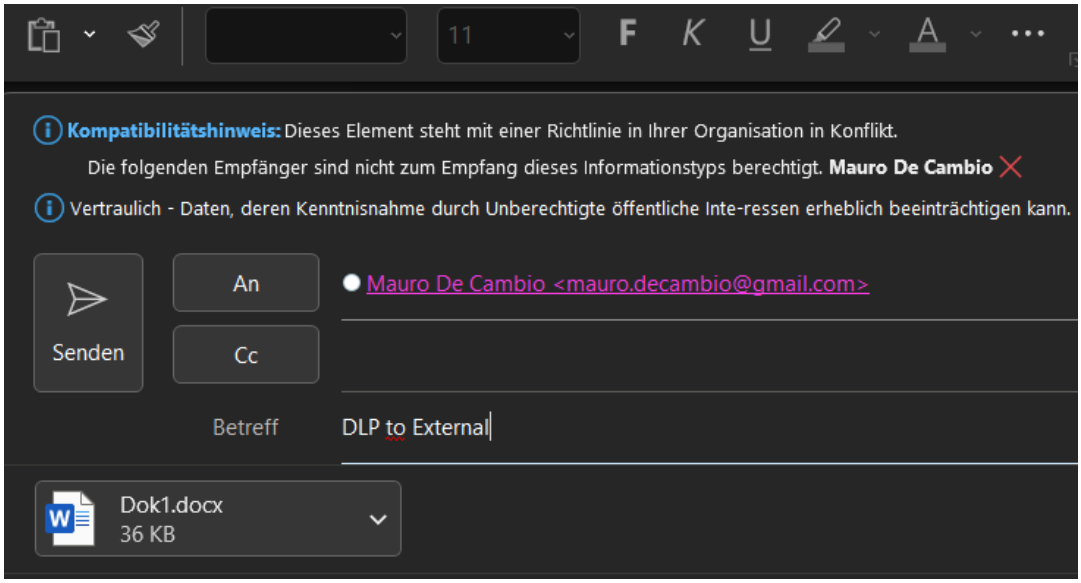

Abbildung 13: Meldung beim Versenden einer vertraulichen Nachricht an einen externen Empfänger

Das direkte Teilen über die Teilen-Funktion in OneDrive und SharePoint weicht davon geringfügig ab. Es kann jeweils ein Link zu der Datei generiert werden und dieser kann auch direkt über die Teilen-Funktion verschickt werden. Versucht jedoch der externe Empfänger, das Dokument über diesen Link abzurufen, erscheint eine Meldung, dass der Zugriff verweigert wird. Dieses Verhalten ist zwar nicht optimal, da der Absender nicht merkt, dass das Versenden der Datei nicht gestattet ist. Da der Datenverlust jedoch trotzdem verhindert wird und das entsprechende Ereignis aufgezeichnet wird, ist dies vernachlässigbar. Möglicherweise kann diese Einstellung jedoch noch verbessert werden.

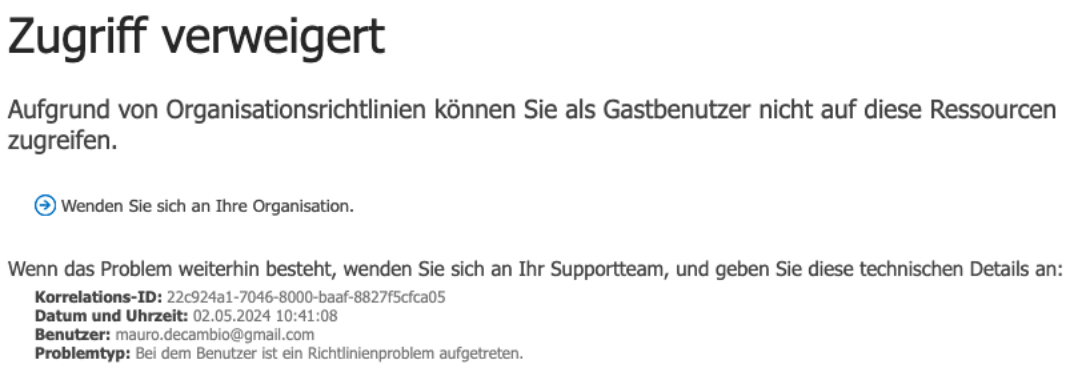

Abbildung 14: Meldung beim Versuch, eine vertrauliche Datei zu öffnen

Kopieren auf externe Speichermedien

Der Datenabfluss über das Versenden oder Hochladen ist jedoch nicht der einzige Risikofaktor. Auch das Kopieren von Dateien auf externe Speichermedien kann mit Microsoft Purview DLP unterbunden werden. In diesem Fall wird eine Richtlinie erstellt, welche das Kopieren der Dateien

auf ein USB-Stick oder Laufwerk, per Bluetooth oder über eine Remote Desktop-Applikation verhindern soll. Zur Überprüfung dieser Richtlinie wird wieder das vertrauliche Dokument Dok1.docx verwendet. Dieses Mal wird versucht, das Dokument aus dem lokalen Speicher auf einen USB-Stick zu kopieren. Auch dieser Versuch wird erfolgreich blockiert. Eine Meldung teilt dem Benutzenden mit, dass dieser Vorgang nicht erlaubt ist. Die Meldung lässt vermuten, dass dies daran liegt, dass das Dokument noch geöffnet ist. Dabei scheint es sich jedoch um einen Übersetzungsfehler zu handeln, da dieses Dokument zum Zeitpunkt des Versuchs nicht geöffnet war.

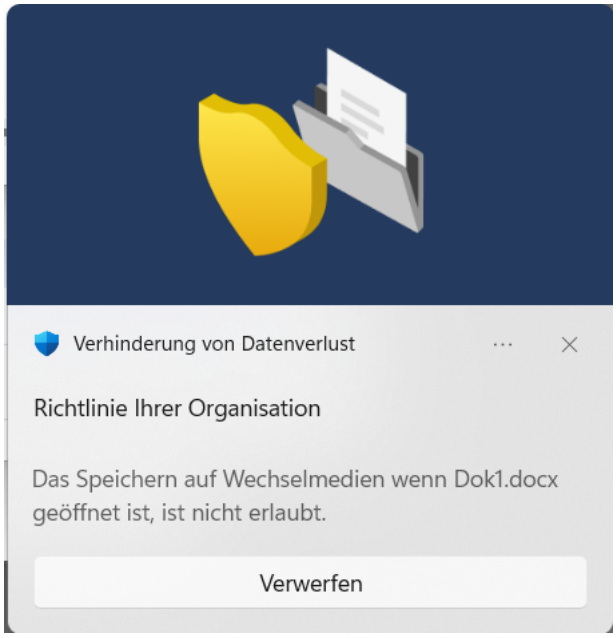

Abbildung 15: Meldung beim versuchten Kopieren auf externes Speichermedium

Upload auf Webseiten und Cloudumgebungen

Ein grosses Risiko für den Datenverlust bieten Webseiten, über welche Daten hochgeladen werden können. So ist es möglich, schnell grosse Mengen an Daten aus dem internen kontrollierten Netzwerk zu extrahieren. Sobald sich die Daten auf einem Speicher befinden, über den die Stadtverwaltung keine Kontrolle hat, können die Daten ungehindert verbreitet und missbraucht werden. Es ist somit wichtig, diesen Abfluss frühzeitig zu unterbinden. Hierfür wird eine entsprechende Richtlinie erstellt. Diese Richtlinie kann kontrollieren, auf welche Zieldomäne eine Datei hochgeladen werden soll. Ist die Zieldomäne in einer Liste mit erlaubten Domänen eingetragen, so wird der Upload erlaubt. Befindet sich die Domäne nicht darauf, wird der Download blockiert.

Jede der bisherigen Richtlinien bietet die Möglichkeit, dass die Mitarbeitenden die Blockade selbst umgehen können. Eine Richtlinie kann mit entsprechender Begründung einmalig ausser Kraft gesetzt werden. Dies wurde zur Veranschaulichung in einem Beispiel umgesetzt.

Das Upload-Ziel ist der Mailservice von Gmail. Es wird versucht, darin eine E-Mail zu erstellen und das vertrauliche Dokument Dok1.docx als Anhang anzufügen. Da dies als Upload gilt, und die Domain mail.google.com nicht explizit erlaubt ist, wird dieser Upload blockiert. Die Blockade kann mit Ausser- Kraft-Setzen umgangen werden. Dieser Vorgang wird, wie jegliches Eingreifen in eine Richtlinie, aufgezeichnet.

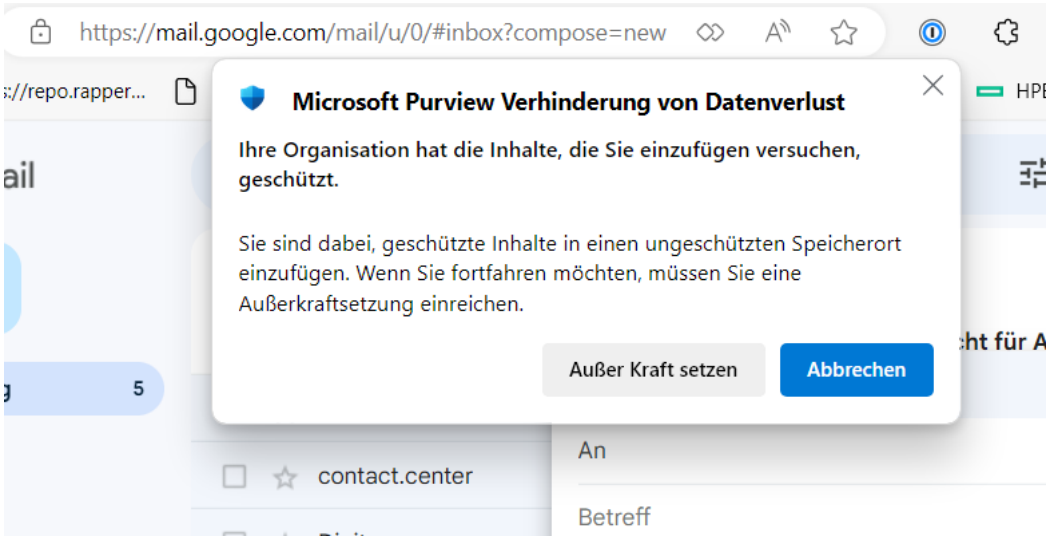

Abbildung 16: Block mit Möglichkeit zur Umgehung beim Upload

Die DLP-Richtlinien sind ein äusserst mächtiges Tool um beabsichtigten so wie unbeabsichtigten Datenabfluss aus der Stadtverwaltung zu erkennen, zu kontrollieren und zu unterdrücken. Bei den erläuterten Richtlinien handelt es sich lediglich um einfache Einstellungen. Es können weitaus feinere Abstimmungen vorgenommen werden, um eine Vielzahl an möglichen Datenlecks zu verhindern. Wie auch bei der Klassifikation gilt es, diese Richtlinien und Mechanismen mit der Zeit weiterzuentwickeln und anzupassen.

6.5.8 Monitoring

Wie bereits erwähnt zeichnet Microsoft Purview eine Vielzahl an verschiedenen Logs auf, welche sogleich in Purview dargestellt werden. Für den PoC sind vor allem zwei Monitoring Tools interessant: der Activity Explorer und die DLP Alerts. Der Activity Explorer überwacht alles, was mit Labels bzw. Klassen zu tun hat und stellt dies dar (Fox, Koenen, Mazzoli, et al., 2024). In der Abbildung 17 ist zu sehen, welche Aktionen für Dateien, welche mit einer Klasse versehen sind, vom Pilotbenutzer an einem Tag ausgeführt wurden.

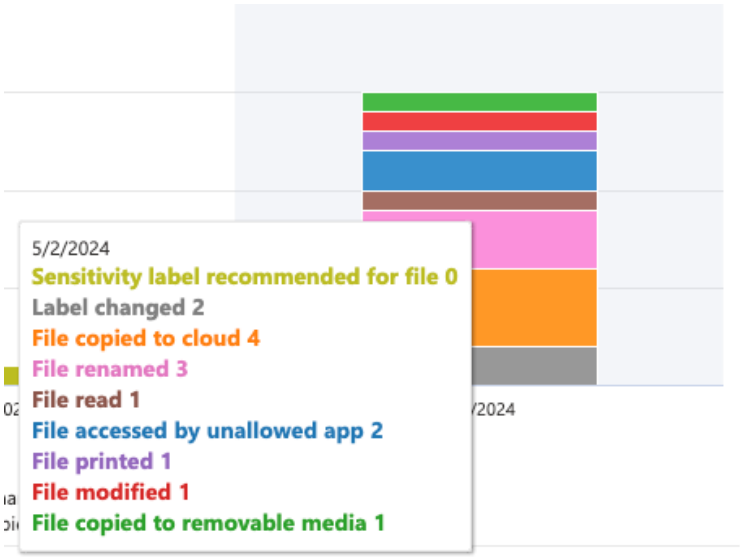

Abbildung 17: Darstellung der Aktionen mit klassifizierten Dateien

Im detaillierten Log können genauere Informationen zu den einzelnen Einträgen aufgerufen werden. Beispielsweise wird der Versuch, die vertrauliche Datei Dok1.docx auf Gmail hochzuladen, aufgezeichnet

Activity	File	Location	User	Happened	Policy	Rule
File copied to cloud	C:\Users\decma\OneDrive - Rapperswil-Jona\DLP\Dok1.docx	Endpoint devices	Mauro.De-Cambio	2024-05-02T 18:28:54.000Z	Block Apps Vertraulich	Restrict Web App

Da dieser Log-Eintrag mit einer Richtlinie und deren Regel übereinstimmt, welche die Auslösung von Alarmen definiert hat, wird gleichzeitig ein Alarm ausgelöst. Dieser Alarm wird im zweiten Tool, den DLP Alerts, aufgezeichnet (Fox, Koenen, & Mazzoli, 2024). DLP Alerts dienen dazu, sich schnell einen Überblick zu verschaffen, ob DLP-Richtlinien bei Mitarbeitenden angewendet wurden und ob mögliche Verstösse vorliegen. Je nach Einstellungen können Alerts per Mail an Administratoren zugestellt werden, welche diese anschliessend bearbeiten sollen.

6.6 Bewertung der Lösung

6.6.1 Sicherheitsanalyse

In diesem Kapitel werden die Funktionen von Microsoft Purview mit den vordefinierten Einschränkungen verglichen, um zu überprüfen, ob die Einschränkungen durchgesetzt werden können.

Das Einteilen der Dokumente in Klassen kann durch Sensitivity Labels gelöst werden. Diese Labels können einzeln definiert werden und haben unterschiedliche Auswirkungen auf die Dateien. Dateien können mit Kopf- und Fusszeilen oder mit Wasserzeichen versehen werden. Interaktionen das Kopieren oder Drucken des Inhalts können aufgrund von einzelnen Accounts oder Benutzergruppen eingeschränkt werden.

Dateien erhalten diese Labels bzw. Klassen entweder durch das manuelle Zuweisen durch die Mitarbeitenden oder durch automatische Mechanismen. Das manuelle Zuweisen funktioniert über ein einfaches Dropdown-Menü in unterstützten Applikationen (Office Applikationen, Adobe Acrobat, usw.). Automatisch lassen sich bei Microsoft Purview Klassen anhand sogenannter Sensitive Information Types (SIT) oder Trainable Classifiern zuteilen. Bei SIT handelt es sich um Muster, welche fix vorgegeben sind und in Dateien vorkommen. Bei Übereinstimmung von Datei und SIT wird eine Klasse zugewiesen. Die Trainable Classifier basieren auf Machine Learning. Welche Algorithmen genau verwendet werden, geht aus der Dokumentation von Microsoft nicht hervor.

Für den PoC wurde ein vorgefertigter SIT für der Erkennung von AHV-Nummern in Dokumenten verwendet. Bei den Tests funktionierte dieser sehr gut und erkannte AHV-Nummern in allen Dokumenten. Für die automatische Erkennung sind jedoch Microsoft 365 E5-Lizenzen nötig. Von diesen besitzt die Stadtverwaltung zurzeit nur einige für Tests. Es ist jedoch geplant, allen Mitarbeitenden eine solche Lizenz zuzuweisen.

Zugriff durch Personen (unabhängig von Klasse immer über Zugriffsrechte geregelt)

> Dies wird einerseits durch das RBAC-Zugriffsmodell in SharePoint geregelt. Zusätzlich dazu, kann der Zugriff durch Verschlüsselung eingeschränkt werden. Dabei können Berechtigungen vordefiniert werden, wie für die Klasse geheim beschrieben. Zusätzlich können diese Zugriffsrechte auch von Mitarbeitenden selbst beim Erstellen des Dokumentes gesetzt werden, was jedoch für diesen PoC nicht benötigt wird.

Öffnen der Datei nur in vorgegebener Software

> Über die DLP-Richtlinien innerhalb von Microsoft Purview kann festgelegt werden, welche Software als sicher eingestuft werden und dadurch erlaubt sind und welche blockiert werden.

Zugriff nur aus bestimmten Ländern (Geoblocking)

Geoblocking wurde in diesem PoC nicht behandelt, da dies bereits flächendeckend bei der Stadtverwaltung im Einsatz ist. Über die Funktion Conditional Access der Microsoft Azure Active Directory wird aufgrund der Source IP der Zugang auf Daten entweder erlaubt oder blockiert.

Einschränkung der zugreifenden Endgeräte

Auch diese Einschränkung ist bereits in Kraft und wird ebenfalls über Conditional Access geregelt. Dabei wird überprüft, ob das Endgerät Mitglied des Mobilen Device Management Tools der Stadtverwaltung ist. Falls es Mitglied ist, wird es durch den Informatikdienst verwaltet und entspricht den Sicherheitsvorschriften. Der Zugriff wird gewährt. Falls das Gerät diesen Mitgliederstatus nicht besitzt oder verliert, wird der Zugang zu den Dateien blockiert.

Verschlüsselung der Datei

Dateien können anhand der Label-Richtlinie verschlüsselt werden. Es kann innerhalb dieser Richtlinien definiert werden, welche Benutzer oder Gruppen, welche Zugriffsrechte auf die Datei besitzen. Um diese Zugriffsrechte umzusetzen, wird die Datei verschlüsselt und nur die berechtigten Personen können diese wieder entschlüsseln.

Upload auf Webseiten und Cloudumgebungen

Diese Einschränkung wird anhand von DLP-Richtlinien durchgesetzt. Aufgrund von Allow- und Blocklisten können Domänen für den Upload von Dateien freigegeben oder blockiert werden. Die Richtlinie kann, wie alle DLP-Richtlinien auf einzelne Klassen unterschiedlich angewendet werden.

Teilen von Daten (Siehe Kapitel 5.3 **)**

Das Teilen von Daten wird ebenfalls durch DLP-Richtlinien blockiert. Es kann zwischen externen und internen sowie zwischen vertrauenswürdigen und nicht vertrauenswürdigen Empfängern unterschieden werden. So können ausführliche Richtlinien zusammengestellt werden, um das Teilen der Daten genau zu steuern.

Kopieren auf externe Speichermedien (USB, Festplatten)

Auch diese Einschränkung wird durch DLP-Richtlinien gesteuert. Grundsätzlich werden alle externen Speichermedien blockiert. Es können jedoch Ausnahmen aufgrund der Hardware-ID eines Speichermediums eingerichtet werden, um Ausnahmen zu erstellen.

Datei ausdrucken

Der Ausdruck einer Datei kann auf zwei Ebenen kontrolliert werden. Es kann per Zugriffsteuerung in den Label-Richtlinien die Möglichkeit, zu drucken, vollständig ausgeblendet werden. Eine feinere Steuerung bieten die DLP-Richtlinien. Hier kann aufgrund des Druckers unterschieden werden, ob der Druck erlaubt ist oder nicht.

Kopieren von Inhalt in die Zwischenablage

Das Kopieren von Inhalten ist über die Label-Richtlinien möglich. Die Datei lässt, sofern festgelegt, kein Kopieren des Inhaltes durch Tastenkombinationen oder Mausinteraktion zu. Des Weiteren können Screenshots unterdrückt werden, indem der Bildschirm beim Versuch der Aufnahme schwarz wird.

7 Fazit

7.1 Zusammenfassung der Ergebnisse

Die Thesis verfolgte das Ziel, ein Konzept zur Datenklassifikation zu entwerfen, welches der Stadtverwaltung Rapperswil-Jona hilft, die Datensicherheit und die Einhaltung des Need-to Know-Prinzips zu verbessern. Die Forschungsfrage diente dabei als Leitfaden durch die Thesis.

Inwiefern trägt die Entwicklung und Implementierung eines Konzepts zur Datenklassifikation dazu bei, das Need-to-Know-Prinzip und die Data Loss Prevention der Stadt Rapperswil-Jona zu verbessern?

Um diese Frage zu beantworten, wurden die bestehenden Sicherheitsinfrastrukturen mit Fokus auf das Berechtigungssystem umfassend analysiert. Durch aus den Analysen hervorgegangene Schwächen des bisherigen Konzepts und die Identifikation von Erfordernissen für eine neue Lösung konnte ein Sicherheitskonzept entwickelt werden, welches aufzeigt, dass anhand der Implementierung von Datenklassifikationssystemen die Datensicherheit massgeblich verbessert werden kann.

Die vorgestellten Methoden zur Datenklassifikation ermöglichen eine präzisere Kontrolle und eine deutlich effektivere Überwachung des Datenzugriffs innerhalb der Stadtverwaltung. Anhand der Klassifizierung von Daten durch Mitarbeitende selbst, aber auch durch automatisierte Systeme, kann eine deutliche Reduktion bestehender Sicherheitsrisiken und potenzieller Datenlecks erzielt werden. Das erarbeitete Konzept unterstützt zudem die Einhaltung des «Need-to-Know»-Prinzips, indem unberechtigte Datenzugriffe weiter reduziert werden können. Die im Konzept vorgeschlagenen Klassen richten sich nach den Empfehlungen des Kantons St.Gallen an die Gemeinden des Kantons. Somit wird sichergestellt, dass Klassen einheitlich verwendet werden.

Während des PoC wurde überprüft, ob eine praktische Umsetzung des zuvor erarbeiteten Konzepts realisierbar und integrierbar ist. Anhand der vorhergehend erarbeiteten Anforderungen an die Datensicherheit wurde eine Lösung ausgearbeitet, welche der Stadtverwaltung hilft, dieses Konzept in die Praxis umzusetzen.

In Anbetracht der Ergebnisse handelt es sich um einen erfolgreichen Proof of Concept. Die Sicherheit der Daten kann durch die Klassifizierung deutlich gesteigert werden, indem fein definierte Richtlinien die Kontrolle über die Daten verbessern. Durch die ausführlichen Monitoring-Möglichkeiten von Microsoft Purview können Verstösse gegen diese Richtlinien oder Versuche,

diese zu umgehen, frühzeitig erkannt werden. Dadurch kann schnell und effizient auf solche Vorfälle reagiert werden. Der Datenverlust wird so verhindert, bevor er passieren kann.

Das Einteilen von Daten in die vier definierten Klassen funktioniert dank der ausgezeichneten Integration in die Office-Applikationen, aber auch in diverse Drittanbieter-Applikationen wie Adobe Acrobat, intuitiv. Dies ist vor allem für die Mitarbeitenden, welche tagtäglich mit diesen Daten arbeiten, sehr angenehm, da keine zusätzliche komplizierte Software eingesetzt werden muss.

Abschliessend kann bestätigt werden, dass durch ein neues Klassifikationskonzept eine deutliche Verbesserung in der Datensicherheit und in dem «Need-to-Know»-Aspekt für die Stadtverwaltung Rapperswil-Jona erreicht werden kann, sofern dieses auch konsequent umgesetzt wird.

7.2 Limitationen und Herausforderungen

Eine Einführung der Dokumentenklassifikation in der Stadtverwaltung bringt einige Herausforderungen mit sich. Ein System, welches fast alle sensitiven Daten der Stadtverwaltung tangiert, kann schnell komplex und unübersichtlich werden. Mit der im PoC vorgeschlagenen Lösung von Microsoft kann dieser Komplexität zwar entgegengewirkt werden, da keine zusätzliche Software nötig ist, jedoch wird die Klassifikationslösung mit der Zeit immer mehr wachsen und an Komplexität gewinnen.

Ein weiterer Faktor sind die Mitarbeitenden, welche jeden Tag Daten generieren und damit arbeiten. Es ist schwierig, die Balance zwischen Benutzerfreundlichkeit und Datensicherheit zu finden. Die Mitarbeitenden wünschen sich auf der einen Seite die Freiheit, so zu arbeiten wie sie es für richtig erachten und es über die Jahre gelernt haben. Andererseits liegt es am Informatikdienst und an der Stadtverwaltung, diese Daten so zu schützen und Datenverlust aktiv zu verhindern, bevor dieser geschieht. Das heisst, dass es zu Einschränkungen in der «Freiheit» der Mitarbeitenden kommen kann und einige nicht so mit den Daten umgehen werden können, wie sie es bisher gewohnt waren. Hier gilt es, einen Mittelweg zu finden.

Wie bereits erwähnt werden die meisten Mitarbeitenden von den Änderungen, welche die Klassifikation mit sich bringt, direkt betroffen sein. Um Komplikationen und Unzufriedenheit während und nach der Einführung zu verhindern, ist es von grosser Bedeutung, die Mitabreitenden in dieser Zeit zu unterstützen. Im Konzept wurden dafür die die Methoden der Schulung und Dokumentation erläutert. In den Schulungen sollen die Mitarbeitenden im Bereich der Datensicherheit sensibilisiert werden. Auch soll Wissen zur Klassifikation im Allgemeinen und zur Software im

Besonderen vermittelt werden. Wissensressourcen wie Dokumentationen helfen dann, dieses Wissen aufrecht zu erhalten, zu vertiefen und zu erweitern.

7.3 Ausblick auf zukünftige Entwicklungen

Eine Klassifikationslösung muss sich in erster Linie an die Bedürfnisse der Stadtverwaltung anpassen. Da sich die Stadtverwaltung jedoch, wie jede Organisation weiterentwickelt und sich damit Organisationsstrukturen und Bedürfnisse verändern, muss auch eine Klassifikationslösung flexibel anpassbar sein. Richtlinien, welche im PoC erarbeitet wurden, werden sich während und nach der Implementierung deutlich verändern. Es werden neue Richtlinien für neue potenzielle Bedrohungen erstellt werden und bestehende verschwinden. So wird die Lösung organisch mit den Anforderungen der Stadtverwaltung wachsen, um diese möglichst lückenlos abzudecken.

Mit den vier Klassengeheim, vertraulich, intern und öffentlich, welche der Kanton vorgibt, können nicht alle Daten abgedeckt werden. In der Stadtverwaltung gibt es kein Verbot für das Speichern und Verwenden privater Daten. Es wäre daher sinnvoll, wenn diese Daten mit einer eigenen Klasse gegenzeichnet werden könnten. So kann verhindert werden, dass private Daten von Einschränkungen betroffen sind, welche sich an die anderen Klassen richten. Damit wird auch verhindert, dass Mitarbeitende diese Daten manuell in eine dieser Klassen einteilen.

Aus den Dokumentationen von Microsoft Purview geht leider nicht hervor, anhand welcher Machine Learning-Algorithmen Daten erkannt werden. Durch die schnelle Entwicklung von Large Language-Modellen (LLM) und Generative AI-Modellen wie ChatGPT ist es mittlerweile möglich, dass der Inhalt von geschriebenem Text und Bildmaterial verstanden und interpretiert werden kann. Diese Technologien haben vor allem in den letzten Jahren einen enormen Aufschwung erlebt. Sie werden in den kommenden Jahren einen bedeutenden Einfluss im Datenmanagement von Organisationen erlangen.

7.4 Kritische Reflexion

Die Datenklassifizierung ist ein Thema, welches viele Unternehmen beschäftigt. Im heutigen Unternehmenszeitalter ist es essenziell, die eigenen Daten zu kennen. Es ist ein wichtiges und spannendes Thema, welches gleichzeitig auch die Stadtverwaltung Rapperswil-Jona vor eine Herausforderung stellt. Die Arbeit ist praktisch ausgerichtet und soll der Stadtverwaltung einen praktischen Nutzen bei der Bewältigung dieser Herausforderung bieten.

Eine Herausforderung vor allem in der Literaturrecherche war, dass sich viel Fachliteratur ausschliesslich auf die Technologien hinter der Datenklassifikation fokussieren. So werden häufig

Modelle miteinander verglichen oder selbst verbessert. Die Operationalisierung einer solchen Lösung wird dagegen eher vernachlässigt.

Eine weitere Schwierigkeit ergab sich durch das Unternehmen, welchem diese Thesis Nutzen bringen soll. Die Stadtverwaltung befindet sich in einer Phase der Reorganisation. Es wurde zu der Zeit, als diese Thesis entstanden ist, aktiv an einer neuen IT-Strategie und an einer neuen Organisation der gesamten städtischen Verwaltung gearbeitet.

Anhang

Quellenverzeichnis

Ashcraft, A., Park, C., Sharkey, K., Coulter, D., Jacobs, M., & Satran, M. (2023, Februar 7). *Access control lists—Win32 apps*. https://learn.microsoft.com/en-us/windows/win32/secauthz/access-control-lists

Ashcraft, A., Sharkey, K., Coulter, D., Jacobs, M., & Satran, M. (2021, Januar 7). *Access Control Entries—Win32 apps*. https://learn.microsoft.com/en-us/windows/win32/secauthz/access-control-entries

Bailey, C., Koenen, K., Cross, K., Mazzoli, R., Buck, A., Henderson, W., & Savell, S. (2024, April 11). *Get started with sensitivity labels*. https://learn.microsoft.com/en-us/purview/get-started-with-sensitivity-labels

Bailey, C., Koenen, K., Henderson, W., Mazzoli, R., & Savell, S. (2023, Dezember 12). *Deploy an information protection solution with Microsoft Purview*. https://learn.microsoft.com/en-us/purview/information-protection-solution

Bailey, C., Koenen, K., Mazzoli, R., Buck, A., Henderson, W., & Savell, S. (2024, März 26). *Automatically apply a sensitivity label in Microsoft 365*. https://learn.microsoft.com/en-us/purview/apply-sensitivity-label-automatically

Bailey, C., Robertson, S., Koenen, K., Buck, A., Henderson, W., Mazzoli, R., & Savell, S. (2024, März 26). *Learn about sensitivity labels*. https://learn.microsoft.com/en-us/purview/sensitivity-labels

Eloff, J. H. P., Holbein, R., & Teufel, S. (1996). Security classification for documents. *Computers & Security*, *15*(1), 55–71. https://doi.org/10.1016/0167-4048(95)00023-2

Enterprise, I. D. G. (1984). *Computerworld*. IDG Enterprise.

Exact Data Matching (EDM). (2024, März 13). Palo Alto. https://docs.paloaltonetworks.com/enterprise-dlp/administration/configure-enterprise-dlp/configure-exact-data-matching

Föck, M., & Fröschle, H.-P. (2011). Schutz der Unternehmensdaten—Data Leakage Protection (DLP). *HMD Praxis der Wirtschaftsinformatik*, *48*(5), 28–34. https://doi.org/10.1007/BF03340621

Fong, A. C. M. (2010). A review of Machine Learning Algorithms for Text-Documents Classification. *Journal of Advances in Information Technology*, *1*(1), 4–20. https://doi.org/10.4304/jait.1.1.1-1

Fox, C., & Koenen, K. (2024a, Februar 7). *About document fingerprinting*. https://learn.microsoft.com/en-us/purview/sit-document-fingerprinting

Fox, C., & Koenen, K. (2024b, Februar 22). *Learn about sensitive information types*. https://learn.microsoft.com/en-us/purview/sit-sensitive-information-type-learn-about

Fox, C., Koenen, K., & Mazzoli, R. (2024, Januar 30). *Get started with data loss prevention alerts*. https://learn.microsoft.com/en-us/purview/dlp-alerts-get-started

Fox, C., Koenen, K., Mazzoli, R., Baldua, T., & Savell, S. (2024, April 19). *Get started with Activity explorer*. https://learn.microsoft.com/en-us/purview/data-classification-activity-explorer

Fox, C., Koenen, K., & Savell, S. (2024, Februar 22). *Learn about exact data match based sensitive information types*. https://learn.microsoft.com/en-us/purview/sit-learn-about-exact-data-match-based-sits

Fox, C., Mazzoli, R., & Koenen, K. (2023, September 26). *Switzerland SSN AHV number entity definition*. https://learn.microsoft.com/en-us/purview/sit-defn-switzerland-ssn-ahv-number

Friedl, J. E. F. (2006). *Mastering Regular Expressions*. O'Reilly Media, Incorporated. https://books.google.ch/books?id=P5UXAwAAQBAJ

Hu, V. C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., & Scarfone, K. (2014). *Guide to Attribute Based Access Control (ABAC) Definition and Considerations* (NIST SP 800-162; S. NIST SP 800-162). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-162

Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent Convolutional Neural Networks for Text Classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*(1). https://doi.org/10.1609/aaai.v29i1.9513

Mazzoli, R., Vukos-Walker, C., & Henderson, W. (2024, April 16). *Learn about Microsoft Purview*. Microsoft Learn. https://learn.microsoft.com/en-us/purview/purview

Microsoft Purview Information Protection setup guide | MIP setup guide. (o. J.). Abgerufen 24. April 2024, von https://setup.cloud.microsoft/purview/information-protection?Q=learndocs

Osborn, S., Sandhu, R., & Munawer, Q. (2000). Configuring role-based access control to enforce mandatory and discretionary access control policies. *ACM Transactions on Information and System Security*, *3*(2), 85–106. https://doi.org/10.1145/354876.354878

Raman, P., Güne Kayacik, H., & Somayaji, A. (2011). Understanding Data Leak Prevention. *6th Annual Symposium on Information Assurance (ASIA '11)*, 36–40.

Rigutini, L., & Maggini, M. (2010). Automatic text processing: Machine learning techniques. *LAP LAMBERT Academic Publishing*.

Rogowski, W. (2013). The right approach to data loss prevention. *Computer Fraud & Security*, *2013*(8), 5–7. https://doi.org/10.1016/S1361-3723(13)70070-8

Ting, S. L., Ip, W. H., & Tsang, A. (2011). Is Naïve Bayes a Good Classifier for Document Classification? *International Journal of Software Engineering and its Applications*, *5*.

Toelle, E. (2021). *Microsoft 365 Compliance: A Practical Guide to Managing Risk*. Apress. https://doi.org/10.1007/978-1-4842-5778-4

Unterstützung für Microsoft Purview Information Protection in Acrobat. (2023, Mai 9). Adobe. https://helpx.adobe.com/content/help/ch/de/enterprise/kb/mpip-support-acrobat.html

Wlosinski, L. (2018, Februar 16). *Data Loss Prevention—Next Steps*. ISACA. https://www.isaca.org/resources/isaca-journal/issues/2018/volume-1/data-loss-prevention-next-steps

Abbildverzeichnis

Tabellenverzeichnis