Onderzoeksrapport

Real-time Compliance Dashboard voor SaaS

Naam student: [Anwar Muradin]

Studentnummer: [500940702]

Bedrijf: SECIAN

Afdeling / Team: [Development APEX dashboard]

Begeleider (bedrijf): [Amar Ritoe]

Begeleider (HvA): [Juriaan Vogelzang]

Datum: [11 mei]

Opleiding

HBO-ICT – [Technische informatica]

Hogeschool van Amsterdam

Stageperiode: [3-Februari – 27-Juni 2025]

Samenvatting

Dit onderzoek richtte zich op het ontwikkelen van een SaaS-dashboard dat realtime inzicht biedt in de compliance status van organisaties die gebruikmaken van Microsoft 365, Google Workspace en AWS. Het platform is ontworpen met als doel om niet-technische bedrijfsgebruikers te ondersteunen bij het naleven van actuele en soms wettelijk verplichte standaarden, zoals ISO27001 en SOC2.

Het ontwikkelde prototype toont hoe authenticatie en gebruikersbeheer veilig kunnen worden geregeld via Microsoft Entra, met duidelijke rolverdelingen tussen adminUsers, regularUsers en companyEntities. Vervolgens werd onderzocht hoe compliance data efficiënt en schaalbaar verwerkt kan worden met behulp van DynamoDB, versleuteling en incrementele updates. Tot slot werd gekeken naar de harmonisatie van logs uit verschillende cloudplatformen, waarbij een gestandaardiseerd datatype en UserID-koppeling centraal staan.

Het resultaat is een ALPHA-release dat compliance data vergelijkt met vooraf gedefinieerde baselines en hierover automatische rapportage mogelijk maakt. De huidige implementatie beperkt zich tot Microsoft 365, met de mogelijkheid tot uitbreiding naar andere cloudplatformen in toekomstige iteraties.

Inhoudsopgave

1. Inleiding en Probleemcontext

1.0 Doel en Relevantie van het Onderzoek

Steeds meer organisaties – van pre-seed start-ups tot gevestigde MKB-bedrijven – draaien hun volledige IT-landschap in de cloud. Diensten als **Microsoft 365**, **Google Workspace** en **Amazon Web Services (AWS)** maken het eenvoudig om werkplekken, data-opslag en applicatie-hosting met één muisklik op te zetten.

Diezelfde laagdrempeligheid creëert echter een **compliance-kloof**. Nieuwe wet- en regelgeving (bijv. AVG/GDPR, ISO 27001, SOC 2) stelt harde eisen aan informatiebeveiliging, terwijl de meeste oprichters en managers geen formele security-achtergrond hebben. Daarbij komt dat de beveiligings- en compliance-controls per cloudplatform sterk **inconsistent** zijn – nomenclatuur, granulariteit en rapportagevormen verschillen. Niet-technische teams kunnen daardoor nauwelijks vaststellen of maatregelen overal correct zijn toegepast. Het handmatig verzamelen van audit-logs, bewijsdocumenten en policy-checks uit elk platform is bovendien tijdrovend, foutgevoelig en kostbaar.

Mijn stagebedrijf, SECIAN, wil die kloof dichten met een SaaS-product: een **geïntegreerd compliance-platform** dat automatisch controles uitvoert, bewijsmateriaal archiveert en één helder dashboard toont. Het onderzoek richt zich op de technische haalbaarheid en architectuurkeuzes om zo'n oplossing schaalbaar, veilig en toch eenvoudig te houden voor een niet-technisch publiek.

2. Hoofd- en Deelvragen

Hoofdvraag

Hoe kan een start-up een schaalbare SaaS-architectuur ontwerpen die realtime compliance-gegevens uit meerdere cloud-platformen consolideert en presenteert aan niet-technische gebruikers?

Deelvragen

1. **Authenticatie & Gebruikersbeheer** – Welke rol- en session-modellen zijn nodig om zowel eindgebruikers als beheerders veilig toegang te geven?
2. **Dataverwerking & Opslag** – Hoe kan compliance-data efficiënt, versleuteld en incrementeel uit diverse APIs worden verwerkt?
3. **Log-normalisatie** – Welke datastructuur maakt logs uit Microsoft 365, Google Workspace en AWS onderling vergelijkbaar?

2.1 Leeswijzer

- Hoofdstuk 2 (Onderzoeksopzet) beschrijft methode, scope en gebruikte tech-stack.
- Hoofdstuk 3 (Authenticatie en Gebruikersbeheer) beantwoordt deelvraag 1.
- Hoofdstuk 4 (Veilige en Schaalbare Verwerking) behandelt deelvraag 2.
- Hoofdstuk 5 (Harmonisatie van Multi-Cloud Logs) behandelt deelvraag 3.
- Hoofdstuk 6 (Conclusies en Aanbevelingen) koppelt de resultaten terug naar de hoofdvraag en schetst vervolgstappen.

2.3 Onderzoeksmethode

Het onderzoek is uitgevoerd tijdens de beginfase van mijn stage bij SECIAN met als doel om een real-time compliance dashboard te ontwerpen en valideren. Hierbij is gebruikgemaakt van een technische en functionele feasibility study die de basis vormde voor de probleemafbakening en de keuze van oplossingsrichtingen.

De methode bestond uit:

1. **Documentanalyse**: Bestuderen van technische documentatie van Microsoft Graph API, Google Workspace API en AWS CloudTrail om vast te stellen welke data beschikbaar is voor export.
2. **Vergelijkende analyse**: Analyse van de compliance-eisen van GDPR, CCPA en ISO27001 en het opstellen van een baseline-model waarmee geëxporteerde gegevens konden worden vergeleken.
3. **SSO-analyse en implementatie**: Onderzoek naar SSO-mogelijkheden per platform, inclusief configuratiehandleidingen en testimplementaties met Azure AD, Google SAML en AWS IAM Identity Center. De implementatie van federated login via AWS IAM Identity Center is gebaseerd op de documentatie van Amazon (Amazon Web Services, n.d.-b).
4. **Dataparsering en mapping**: Uitwerken van een parseringsstrategie waarmee logbestanden en metadata uit verschillende bronnen konden worden geharmoniseerd naar één intern datatype. Hierbij zijn de ruwe logs omgezet naar gedefinieerde auditcategorieën (zoals: toegang, encryptie, configuratie).
5. **Toolselectie en prototypebouw**: Eerste verkenning van visualisatietools en dashboarding-frameworks (zoals Grafana en Recharts) met als doel het bouwen van een klikbaar prototype.

De aanpak was iteratief van aard: op basis van bevindingen uit de feasibility study zijn vervolgstappen en technische keuzes voortdurend bijgesteld.

3. Authenticatie en Gebruikersbeheer

3.0 Onderzoeksaanpak

Voor deze deelvraag heb ik mij gericht op gebruikersbeheer en authenticatie via Microsoft Entra. Ik heb eerst de documentatie van Microsoft Graph en Entra ID doorgenomen om te bepalen welke API-endpoints nodig waren voor sessiebeheer en roltoewijzing. Vervolgens heb ik een testomgeving opgezet waarin ik meerdere gebruikersrollen (adminUser, regularUser) heb gesimuleerd. Ik heb geëxperimenteerd met tokenopslag (via cookies en localStorage) en deze gekoppeld aan unieke UserIDs in DynamoDB. Ook heb ik validatietests uitgevoerd waarbij tokens ongeldig werden gemaakt om sessieveiligheid te garanderen.

3.1 Rolgebaseerd Gebruikersbeheer

Binnen het SaaS-platform wordt onderscheid gemaakt tussen drie entiteitstypen: **adminUsers**, **regularUsers** en **companyEntities**. Een **adminUser** registreert een nieuwe organisatie, die vervolgens wordt opgeslagen als een **companyEntity**. Deze **admin** bepaalt welke gegevens en functionaliteiten uit de **companyEntity** toegankelijk zijn voor **regularUsers**, waarmee gedetailleerde toegangscontrole mogelijk is.

Deze structuur maakt het mogelijk om toegangsrechten te beheren op basis van rollen en verantwoordelijkheden. **adminUsers** kunnen voor elk onderdeel van de compliance-data aangeven of reguliere gebruikers dit mogen inzien, bewerken of enkel lezen. De koppeling tussen gebruikers en hun organisatie gebeurt op basis van een intern **UserID** dat uniek is binnen de database. Dit **UserID** wordt als referentie gebruikt bij alle dataverzoeken en sessies, waardoor veilige en consistente koppeling gegarandeerd is.

3.2 Implementatie van Microsoft Entra-authenticatie

Bij **login** en **logout** worden sessie-tokens vernieuwd. Op het clientapparaat wordt tijdelijk een token opgeslagen, bijvoorbeeld in een cookie of via **localStorage**. In de backend wordt dit token gekoppeld aan een intern **UserID**, versleuteld opgeslagen in de database en verwijderd zodra de sessie veilig kan worden beëindigd. De prioriteit ligt echter altijd bij het onmiddellijk verwijderen van het lokale token, gezien het verhoogde risico bij lokale opslag. De initiële authenticatie en autorisatie verlopen via Microsoft Entra (Microsoft, n.d.-b); zodra het sessie-token is verkregen, kan de feitelijke data real-time worden opgehaald via de Microsoft Graph API (Microsoft, n.d.-a). De inloggegevens zijn hiervoor slechts éénmalig nodig.

4. Verwerking en Opslag van Compliance-Data

4.0 Onderzoeksaanpak

Voor deze deelvraag heb ik DynamoDB ingericht als centrale opslag voor compliance data. Ik heb documentatie van de Graph API bestudeerd om te bepalen welke logs automatisch opgehaald konden worden. Op basis daarvan heb ik een script ontwikkeld dat incrementeel data ophaalt en alleen gewijzigde records bijwerkt. Ik heb DynamoDB zo geconfigureerd dat opslag versleuteld plaatsvindt en gecontroleerd of dit voldoet aan ISO27001/SOC2-eisen. Daarnaast heb ik met testdata geëxperimenteerd om de impact op verwerkingstijd en API-belasting te meten bij verschillende datasetgroottes.

4.1 Veilige en Schaalbare Verwerking van Compliance-Data

Binnen de ALPHA-fase van het project lag de nadruk op het ophalen, verwerken en opslaan van compliance data afkomstig uit Microsoft 365. Deze data bevat onder andere auditlogs, gebruikersactiviteiten en metadata zoals bestandsnamen of toegangsrechten. De verwerking van deze data vereist een infrastructuur die zowel schaalbaar als veilig is, met mogelijkheden voor incrementele verwerking en versleutelde opslag. In deze deelvraag wordt ingegaan op hoe deze compliance data technisch wordt verwerkt na authenticatie.

4.2 Versleutelde Opslag in DynamoDB

De gegevensopslag is ondergebracht in een AWS-DynamoDB omgeving die gebruik maakt van encryptie (Amazon Web Services, n.d.-c). Hierdoor is de database in lijn met ISO27001 en SOC2-eisen (International Organization for Standardization, 2022; American Institute of Certified Public Accountants, 2017). De opslag betreft zowel metadata (tokens, sessies) als dynamisch opgehaalde compliance data, zoals auditlogs en gebruikersactiviteiten uit Microsoft 365.

4.3 Incrementiële Dataverwerking

Het platform verwerkt alleen relevante data en werkt met incrementele updates. In plaats van grote datasets steeds opnieuw te laden, worden alleen gewijzigde waarden aangepast. Dit verlaagt de druk op de infrastructuur, verkort verwerkingstijden en voorkomt dubbele opslag. Voor het ophalen van data uit Microsoft Graph wordt gebruik gemaakt van selectieve API-aanroepen die beperkt zijn tot de noodzakelijke endpoints (Microsoft, n.d.-a).

4.4 Architectuur voor Toekomstige Schaalbaarheid

Hoewel het huidige proof-of-concept (ALPHA-fase) zich beperkt tot Microsoft-integraties, is de architectuur opgezet met het oog op uitbreidbaarheid naar Google Workspace en AWS. Door gebruik te maken van een gestandaardiseerde interne API en het unified datatype-concept (zie Deelvraag 3), kan nieuwe data eenvoudig worden toegevoegd en verwerkt met minimale aanpassingen aan de infrastructuur.

5. Harmonisatie van Multi-Cloud Logs

Een van de grootste technische uitdagingen bij het combineren van compliance- of auditlogs uit verschillende cloudplatformen zoals Microsoft 365, Google Workspace en AWS is het verschil in logformaten en datastructuren. Elke leverancier hanteert zijn eigen terminologie, JSON-schema's en veldnamen, waardoor het lastig is om deze logs op een uniforme en beheersbare manier te verwerken. Dit zal in het automatisch uitlezen voor problemen zorgen.

5.0 Onderzoeksaanpak

Bij de harmonisatie van logs heb ik de ruwe output van Microsoft 365 en Google Workspace verzameld via hun respectievelijke API's (Microsoft, n.d.-a; Google, n.d.). Ik heb een parser ontworpen die logregels vertaalt naar een intern unified datatype met vaste categorieën zoals toegang, encryptie en configuratie. Daarbij heb ik geëxperimenteerd met edge cases zoals ontbrekende veldnamen of foutmeldingen. Deze parsed logs heb ik vervolgens getest in een visualisatieprototype (met filtering binnen DynamoDB), om te controleren of entries correct gekoppeld bleven aan hun bijbehorende UserID's. Dit bevestigde dat het unified datatype in de praktijk toepasbaar was. Ook heb ik onderzocht hoe meerdere gebruikers op één subwallet kunnen worden gemapt, en hoe die toegang correct te loggen is binnen het systeem, inclusief differentiatie van rechten en traceerbaarheid.

5.1 Noodzaak van een Unified Datatype

Om dit probleem op te lossen, wordt binnen het platform gebruikgemaakt van een gestandaardiseerd "unified datatype" dat alle externe logs vertaalt naar een gemeenschappelijk intern formaat. Deze uniformiteit maakt het mogelijk om de data centraal te parsen, visualiseren en analyseren zonder dat per bron afwijkende logica nodig is.

5.2 UserID-Koppeling voor Dataconsistentie

De kern van deze uniforme dataverwerking ligt in de wijze waarop gebruikers worden gekoppeld. Elk platformaccount (bijvoorbeeld van Microsoft Entra of Google Workspace) wordt gekoppeld aan een intern UserID dat onafhankelijk is van de dashboard-authenticatie. Hierdoor blijven gevoelige inloggegevens gescheiden van operationele metadata, wat bijdraagt aan de veiligheid van het systeem.

Een eenvoudige manier om deze structuur te begrijpen is via een "wallet-in-wallet"-constructie — een metafoor die in dit verslag wordt gehanteerd om de logische

scheiding tussen hoofdgebruikers en gekoppelde platformaccounts te verduidelijken. Elke gebruiker op het platform heeft één centrale identiteit (de 'wallet'), waarin meerdere sub-identiteiten ('subwallets') worden beheerd. Deze subwallets vertegenwoordigen de verbonden externe accounts zoals Microsoft Entra, Google Workspace en AWS. Zo blijft het beheer centraal en overzichtelijk, terwijl data per platform toch apart verwerkt en gekoppeld kan worden.

Één subwallet kan bovendien door meerdere gebruikers benaderd worden. Denk bijvoorbeeld aan twee werknemers (regularUsers) die beide toegang hebben tot hetzelfde Microsoft Entra-beheeraccount. In dit geval kan een beheerder (adminUser) bepalen welke aspecten en datatoegang beschikbaar zijn voor welke reguliere gebruiker, wat het beheer van rechten en verantwoordelijkheden op detailniveau mogelijk maakt.

5.3 Fijnmazige Toegangscontrole en Logging

Zodra een regularUser is ingelogd, kan het platform — op basis van het gekoppelde UserID — alle bijbehorende data en rechten ophalen via de interne API. Dit maakt het mogelijk om toegangsrechten dynamisch toe te passen en per gebruiker of rol te beperken tot alleen de relevante informatie.

Deze controle op detailniveau over toegangsrechten is een bewuste keuze geweest vanuit compliance- en securityoverwegingen, waarbij de admin user optreedt als controlepunt om dataminimalisatie, rolgebaseerde toegang en traceerbaarheid te waarborgen. Daarnaast heeft de admin user inzage in de toegangsgeschiedenis van de reguliere gebruikers. Dit is essentieel omdat zij gevoelige bedrijfsdata kunnen benaderen of wijzigen. Door logging op gebruikersniveau beschikbaar te maken voor de beheerder, kan de integriteit van het systeem beter worden gegarandeerd en worden verdachte activiteiten tijdig opgemerkt.

6. Conclusie

Het doel van dit project was om een SaaS-oplossing te ontwikkelen die niet-technische bedrijfseigenaren en managers ondersteunt bij het naleven van actuele — en soms wettelijk verplichte — normen op het gebied van cybersecurity en gegevensbescherming. Door compliance-informatie uit platforms zoals Microsoft 365, Google Workspace en AWS automatisch te verzamelen en overzichtelijk weer te geven in één dashboard, wordt het naleven van standaarden toegankelijker gemaakt voor eindgebruikers zonder technische achtergrond. Daarbij kunnen veel onderdelen van regulatie automatisch worden geëvalueerd en getoond.

Deelvraag 1 liet zien hoe authenticatie en gebruikersbeheer is opgezet. Door onderscheid te maken tussen adminUsers, regularUsers en companyEntities ontstaat een toegangsstructuur die schaalbaar en beheersbaar is. De koppeling via unieke UserIDs maakt veilige en consistente sessies mogelijk.

Deelvraag 2 richtte zich op de technische verwerking van compliance data. Door gebruik te maken van DynamoDB-opslag, versleuteling, incrementele updates en API-selectie wordt data efficiënt en veilig verwerkt. De huidige (ALPHA-release) implementatie beperkt zich tot Microsoft, maar is voorbereid op uitbreiding naar andere platformen.

Deelvraag 3 behandelde de normalisatie van logs afkomstig van verschillende cloudproviders. Door de inzet van een unified datatype en een UserID-koppelingsstrategie kunnen logs uniform verwerkt worden, ongeacht bronformaat of structuur. AdminUsers houden controle over wie wat mag zien en bewerken, inclusief logging op gebruikersniveau.

Het project toont aan dat het technisch haalbaar is om een real-time compliance dashboard te bouwen dat aansluit op meerdere grote cloudplatformen. In een volgende fase is het aan te raden om ook Google Workspace en AWS volledig te integreren. Daarnaast zou een gebruikersinterface kunnen worden ontwikkeld die concreet uitlegt welk compliance-risico automatisch is gedetecteerd en welke actie wordt aanbevolen.

Literatuurlijst

- American Institute of Certified Public Accountants. (2017). *Trust services criteria for security, availability, processing integrity, confidentiality, and privacy (SOC 2)*. https://www.aicpa.org/resources/article/trust-services-criteria

- International Organization for Standardization. (2022). *ISO/IEC 27001:2022 – Information security, cybersecurity and privacy protection – Information security management systems – Requirements*. https://www.iso.org/standard/27001
- Microsoft. (n.d.-a). *Overview of Microsoft Graph*. Microsoft Learn. https://learn.microsoft.com/en-us/graph/overview
- Microsoft. (n.d.-b). *What is Microsoft Entra ID?* https://learn.microsoft.com/en-us/entra/fundamentals/whatis
- Google. (n.d.). *Admin SDK overview*. Google Workspace Admin Help. https://developers.google.com/admin-sdk
- Amazon Web Services. (n.d.-a). *What is AWS CloudTrail?* https://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-user-guide.html#:~:text=AWS%20CloudTrail%20is%20an%20AWS,recorded%20as%20events%20in%20CloudTrail.
- Amazon Web Services. (n.d.-b). *Getting started with IAM Identity Center*. https://docs.aws.amazon.com/singlesignon/latest/userguide/what-is.html
- Amazon Web Services. (n.d.-c). *Amazon DynamoDB encryption at rest*. https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/EncryptionAtRest.html

De inhoud van het onderzoek is gebaseerd op eigen stagewerkzaamheden, technische documentatie en projectervaring. AI is uitsluitend ingezet als schrijfhulp onder verantwoordelijkheid van de student.