19

SAI UPDATE

DIGITAAL TIJDSCHRIFT VOOR SAI.BE-LEDEN | SEPTEMBER 2024 | SAI

HOE DIVERS IS JOUW IT-TEAM?

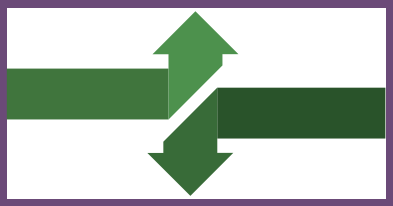
VAN GENDER- TOT LEEFTIJDSDISCRIMINATIE: DE CIJFERS

pag. 10

En verder:

KENNIS OVER GEN AI | BECODE IS VEEL MEER DAN CODE | DE BELANGRIJKSTE REGEL VOOR WACHTWOORDEN
DE COMEBACK VAN TAPEOPSLAG | 50 JAAR COMPUTERWETENSCHAPPEN | DATAVISUALISATIE | DE HACK VAN DE EEUW

CASE VDAB
CONTAINERS OF VIRTUELE SERVERS?
pag. 7



VAN DATA TOT INFRA
DEZE IT-SKILLS ZITTEN IN DE LIFT
pag. 21

CASE SMALS HERGEBRUIK SOFTWARECOMPONENTEN
pag. 28

EDITO

WAT MOGEN WE NOG VERTROUWEN IN DEZE WERELD VAN VERBONDEN TOESTELLEN?

Ik had een mooi edito geschreven, maar de gebeurtenissen van midden juli (CrowdStrike BSOD) en midden september (pagerontploffingen Libanon) noopten me om dit te herschrijven. Steeds meer toestellen zijn verbonden met het internet, niet alleen de traditionele toestellen met browsers, zoals desktops, laptops, tablets en smartphones. Ook wasmachines, koelkasten of inlichtingenborden in stations en luchthavens. Zelfs bedrijfskritische zaken, zoals thermostaten, pacemakers en kerncentrales. Maar deze verbonden toestellen vormen zeer aantrekkelijke doelwitten voor cybercriminelen en mensen met kwade bedoelingen.

De gevaren van verbonden apparaten kunnen zeer ernstige gevolgen hebben, zoals we konden observeren. Fabrikanten kunnen fouten maken in de (update van) producten, waardoor alle geconnecteerde toestellen stilvallen of gewoon niet meer doen wat ze moeten doen, met gigantische cascade-effecten en economische schade. Cybercriminelen kunnen persoonlijke en gevoelige gegevens stelen (zoals wachtwoorden, bankgegevens en andere privé-informatie) of kunnen ransomware installeren op verbonden toestellen, waardoor gebruikers gedwongen worden om losgeld te betalen om weer toegang te krijgen tot hun gegevens.

Geconnecteerde toestellen kunnen gebruikt worden om Distributed Denial of Service (DDoS)-aanvallen uit te voeren, waarbij netwerken overspoeld worden en onbereikbaar worden. Of de toestellen kunnen toegang geven tot camera's en microfoons, waardoor ze gebruikers kunnen bespioneren en zelfs tot ontploffing gebracht kunnen worden.

In zijn boek *If It's Smart, It's Vulnerable* van augustus 2022 waarschuwde Mikko Hypponen al dat alles wat met het internet verbonden is een doelwit wordt. Deze kwetsbaarheid werd aangetoond door de recente massasabotage van beepers in wat experts beschouwen als een zeldzame en uitgebreide aanval. Beepers werden vanop afstand tot ontploffing gebracht, waardoor duizenden mensen ernstig gewond raakten. De aanval dient als een ijzingwekkende herinnering: hoe meer we met elkaar in verbinding staan, hoe meer we worden blootgesteld.

In september 2018 maakte Bruce Schneier in zijn boek *Click here to kill everybody* zich al zorgen over de risico's van verbonden toestellen, omdat de groei te snel gaat en veel minder veilig is dan we zouden willen. Maar bedrijven zijn helaas meestal gericht op winstmaximalisatie. Een gebrekkige beveiliging vormt dikwijls geen bedreiging voor de winst omdat klanten meestal voor de gevolgen opdraaien. En overheden werken soms cyberveiligheid tegen. Ze proberen beveiligingsprotocollen zwak te houden en bepaalde kwetsbaarheden geheim te houden, wat het werk van gerecht en inlichtingendiensten vergemakkelijkt. Cybercriminelen zullen echter altijd hun weg vinden. Het houdt hen niet tegen, maar maakt de wereld een stuk minder veilig voor gewone individuen.

Plots blijken de Europese NIS2-richtlijn (vanaf 18 oktober 2024 in volle werking) en de toekomstige Europese CRA-verordening (in 2027 in volle werking) nog zo slecht niet. Ze waarborgen een hoog gemeenschappelijk veiligheidsniveau binnen de EU via een controlekader om de cyberveiligheid te verbeteren en de risico's van cyberaanvallen te verminderen, ook van geconnecteerde toestellen. Bedrijven en hun toeleveranciers moeten maatregelen nemen om cyberrisico's te beheersen. De NIS2-richtlijn is essentieel om de integriteit en vertrouwelijkheid van onze digitale infrastructuur – en het vertrouwen hierin – te waarborgen. Europa kan daadwerkelijk iets doen door alle bedrijven wettelijk te dwingen om bekend te maken of hun producten bepaalde beveiligingsfuncties hebben en transparant te zijn over datalekken en andere cyberincidenten. Wederom zullen andere landen in andere continenten later volgen.

Een van de grootste risico's voor organisaties en overheden om hun cyberbeleid goed te krijgen, is dat beleidsmakers geen technologische inzichten hebben. Er is nog altijd een enorme behoefte aan mensen met een technische achtergrond die geïnteresseerd zijn in cyberveiligheid en zich ermee gaan bemoeien. Laten we samen nu de nodige acties ondernemen.

Marc Vael
Voorzitter raad van bestuur SAI.BE

INHOUD

"

LAURA 2014! IS GEEN VEILIG WACHTWOORD.

REINAERT VAN DE CRUYS
ETHICAL HACKER
OP PAG 26 ÉN OP HET SAI.BE-WEBINAR

GROTE VERSCHILLEN IN KENNIS OVER GEN AI TUSSEN SECTOREN

Hoe goed kennen medewerkers de mogelijkheden en reikwijdte van Generatieve AI? Dat verschilt weleens tussen sectoren, zo blijkt.

De ideale investering in Generatieve AI (Gen AI) biedt duidelijke mogelijkheden voor efficiëntie en een betere klantervaring, maar veel organisaties melden nog hiaten die een succesvolle uitrol in de weg staan. Dat blijkt uit recent onderzoek van SAS bij 1.600 beslissingsnemers rond AI in Amerika, APC en Europa, waaronder ook België.

Voor veel organisaties is het nog een uitdaging om de kloof van technologie naar uitrol te overbruggen. Dat begint al bij het correct inschatten van de mogelijkheden en impact op hun terrein, ook naar de toekomst toe. Negen op de tien senior beslissers (93%) uit het onderzoek geven toe dat ze Gen AI of de potentiële impact ervan op bedrijfsprocessen (nog) niet volledig begrijpen.

En dan de link met de concrete business case. Bijna de helft (47%) ondervindt problemen bij de overgang van concept naar praktisch gebruik van Gen AI. Meer dan een derde (37%) voorziet problemen om te bewijzen dat Gen AI een sterke ROI biedt of heeft ondervonden dat dit moeilijk valt aan te tonen of te bewijzen.

CIO EN CTO

Ook bij de IT-managementprofielen uit het onderzoek is Gen AI op bepaalde vlakken vaak nog onbekend terrein. **Minder dan de helft (45%) van de CIO's en ongeveer een derde (36%) van de CTO's beschouwt zichzelf als zeer bekend met de adoptie van Gen AI.** Uit het onderzoek blijkt ook dat vier op de tien organisaties (39%) geen Gen AI-gebruiksbeleid hebben voor hun personeel.

Opvallend is ook het verschil tussen de sectoren. Respondenten in manufacturing (waar 61% een goed of volledig begrip aanhaalt), life sciences (55%) en telecom (55%) rapporteren de hoogste niveaus van persoonlijke kennis over Gen AI. Respondenten in de publieke sector hebben veel minder vertrouwen in hun begrip van Gen AI, want daar stelt slechts 38% het concept goed of volledig te begrijpen.

KENNIS VAN GEN AI PER SECTOR (van hoog naar laag)

Manufacturing: **61%**
Life sciences: **55%**
Telecom: **55%**
...
Publieke sector: **38%**

Bron: SAS, 2024

» *Schrijf je als SAI.BE-lid gratis in voor onze avondconferentie 'TechnoVision 2025: IT-trends that will become mainstream in 2025'. Die vindt plaats op 12 december in Van der Valk Brussels Airport Hotel. Meer info vind je **hier**.*

WAAROM WORDEN ZO VEEL CRIMINELE NETWERKEN GEKRAAKT?

Europol kondigde recent de ontmanteling aan van het versleutelde communicatieplatform Ghost, dat voornamelijk gebruikt werd door criminele netwerken. Ghost was een relatief klein netwerk, maar wel het zoveelste dat door de knieën gaat.

Na een drie jaar durend internationaal onderzoek werden naar aanleiding van het ontmantelen van Ghost tientallen arrestaties verricht in onder meer Australië, Ierland en Italië. De actie volgt eerdere successen, zoals de inbeslagname van EncroChat en Sky ECC in 2020 en FBI's Anom-stingoperatie in 2021.

Ook Ghost, een end-to-end encrypted service, werd beschouwd als een veilige haven voor de georganiseerde misdaad. "Dit was een levenslijn voor serieuze en georganiseerde criminaliteit", aldus Jean-Philippe Lecouffe van Europol. Ondanks de geavanceerde technologie van het platform, slaagde de politie erin om de communicatie van gebruikers te onderscheppen en toegang te krijgen tot essentiële data. De inzet van technische middelen, zoals de expertise van het Project Overclock en de samenwerking tussen verschillende landen, bleek cruciaal in de operatie.

Waarom is Ghost slechts het zoveelste platform dat wordt gekraakt? **Criminele netwerken maken steeds gebruik van nieuwe, kleinere versleutelde netwerken om hun communicatie te beschermen.** Lecouffe benadrukte echter dat Europol inmiddels de go-to organisatie is geworden om dergelijke netwerken aan te pakken: "Dit is niet onze eerste ontmanteling en zeker niet de laatste."

Het gebruik van dergelijke platformen blijft een uitdaging voor de opsporingsdiensten, maar internationale samenwerking en technologische innovaties maken het steeds moeilijker voor criminelen om hun activiteiten verborgen te houden. Volgens Lecouffe zal geen enkel platform ongrijpbaar blijven.

MEER GEFRAGMENTEERD

Europol zegt dat de markt steeds meer gefragmenteerd is geraakt door de acties tegen EncroChat, Sky ECC, Ennetcom, Ironchat en An0m. Na de operaties tegen die grote spelers zijn er andere diensten gesloten en weer kleinere opgekomen, wat heeft geleid tot versnippering. Overigens doen internationale politiediensten en regeringen moeite om alle chatdiensten ertoe te brengen de inhoud van communicatie voor overheidsdiensten leesbaar te maken, dus ook die van WhatsApp, Signal en Apple. Zo is deze zomer in Frankrijk de eigenaar van Telegram gearresteerd.

» *Lees ook de boekbespreking over 'De hack van de eeuw', verder in deze editie van SAI Update.*

» *Volg op 15 oktober als SAI.BE-lid gratis het SAI.BE-webinar 'Privacy in Practice with Smart Pseudonymization'. Meer info vind je **hier**.*

ÉÉN SEIZOEN *HOUSE OF THE DRAGON* PER DAG: DATAVERBRUIK BEREIKT NIEUWE MIJLPAAL

Het mobiele dataverbruik piekt als nooit tevoren en ligt vandaag in de Verenigde Staten dubbel zo hoog als twee jaar geleden. Ook in ons land stijgt het verbruik aanzienlijk.

In 2023 verbruikten Amerikanen meer dan 100 biljoen MB aan draadloze data, wat genoeg is om elk huishouden in de Verenigde Staten elke dag het volledige eerste seizoen van *House of the Dragon* te laten streamen. Dat blijkt uit cijfers van CTIA, de vereniging van de mobiele industrie in de Verenigde Staten.

Dit indrukwekkende dataverbruik vertegenwoordigt een stijging van 26 biljoen MB, een groei van 36% ten opzichte van 2022. Het is bijna een verdubbeling tegenover twee jaar geleden. Het markeert de grootste jaarlijkse stijging van draadloze data ooit.

Deze explosieve groei weerspiegelt de steeds grotere rol van 5G-technologie in het dagelijks leven, met toepassingen zoals thuisbreedband, *connected healthcare* en autonoom transport. Van 2010 tot 2018 gebruikten Amerikanen samen minder data dan in 2023 alleen al. Uit cijfers van Ericsson blijkt dat **het datagebruik in de Verenigde Staten tegen 2029 meer dan zal verdrievoudigen**. De infrastructuur rondom 5G-netwerken wordt steeds belangrijker om aan deze vraag te kunnen voldoen.

EN IN BELGIË?

De volumes liggen bij ons lager, maar de groei is evenzeer significant. In zijn jaarlijkse rapport van deze zomer stelt ook het Belgisch Instituut voor postdiensten en telecommunicatie (BIPT) een aanzienlijke stijging van het mobiele dataverkeer vast. Het mobiele dataverkeer steeg in drie jaar tijd met bijna 130 procent.

Het gemiddelde maandelijks dataverbruik per actieve datasimkaart steeg in ons land in 2023 naar 7,5 GB, vergeleken met 4,9 GB eind 2021 en 1,9 GB eind 2018. Ook de uitrol van 5G versnelt.

COMPUTERWETENSCHAPPEN BLAAST VIJFTIG KAARSJES UIT

Vijftig jaar geleden studeerden aan de KU Leuven de eerste studenten af met een masterdiploma computerwetenschappen. De opleiding was toen uniek binnen Europa.

De eerste lichting telde zeventien pioniers die in 1974 als eersten een masterdiploma computerwetenschappen behaalden. De opleiding werd gedoceerd door professoren van de afdeling toegepaste wiskunde en programmatie en werd georganiseerd door de faculteit toegepaste wetenschappen. De belangstelling voor het vakgebied groeide evenwel snel. Enkele jaren later, in 1981, werd het departement computerwetenschappen opgericht. Al was het toen nog met een eerder beperkte bezetting. **Zes professoren en 21 onderzoekers waren toen aan boord.**

Vandaag telt het departement bijna 400 medewerkers, verspreid over 8 campussen in Vlaanderen. Naarmate het aantal onderwijsprogramma's groeide, nam ook het aantal studenten toe. Momenteel volgen meer dan 2.000 studenten opleidingen waarin het departement computerwetenschappen een centrale rol speelt.

Het Leuvense departement computerwetenschappen is nog behoorlijk jeugdig in vergelijking met de KU Leuven zelf. Die wordt in 2025 maar liefst zeshonderd jaar oud.

ARBEIDSBEMIDDELAAR BOUWT VIRTUELE SERVERS AF EN KIEST CONTAINERTECHNOLOGIE

VDAB TREKT EEN BLIK CONTAINERS OPEN

In zijn streven naar een meer datagedreven organisatie trekt de Vlaamse arbeidsbemiddelaar VDAB resoluut de kaart van containertechnologie. Het aantal virtuele machines binnen de organisatie wordt afgebouwd. Een besparing van een tiental miljoen euro op vijf jaar tijd, al zijn er ook andere voordelen.

Officieel is VDAB dan wel een arbeidsbemiddelaar, zelf vinden ze de term loopbaanregisseur meer gepast. Hun beleid rond samenwerking met publieke en private partners moet de werkzaamheidsgraad in Vlaanderen verder opkrikken.

Een centrale datagedreven en resultaatgerichte aanpak staat hierbij centraal, benadrukt Steven Van Peteghem, directeur I&T operations bij VDAB. “Als snelle en wendbare IT-organisatie willen wij ook vlot en snel kunnen inspelen op vragen en uitdagingen”, stelt hij. Door de flinke versnelling in digitalisering bij VDAB, zijn de IT-budgetten op zes à zeven jaar tijd verdubbeld. De I&T-organisatie telt vandaag ongeveer 470 medewerkers. Zowel de *software factory* als architectuur en innovatie kregen meer middelen.

HAAL DE CONTAINERS BINNEN

Het infrastructuurproject rond containers is een belangrijk project binnen de totale aanpak die de I&T-maturiteit bij VDAB verder moet verbeteren**. "We beschikten over 4.500 virtuele servers. Maar vandaag zijn er met containers technische oplossingen die meer geschikt zijn"**, stelt Steven Van Peteghem.

Welke elementen maakten de businesscase rond containers interessant? In eerste instantie is het een puur technisch verhaal. "Het gebruik van systeemresources verloopt optimaler doordat er minder overhead op de OS is", oppert Van Peteghem. "En in combinatie met onze Infra-as-Code-aanpak is er ook een verhoogde stabiliteit en betere recovery van de omgeving."

DE DEVOPS-CASE

De technische upgrade is maar één kant van de zaak, want de businesscase is meer dan dat. "Met name in de DevOps-operaties komt het containerproject van pas", benadrukt de directeur I&T operations. Het project bevordert de samenwerking tussen operaties en ontwikkeling. "We mikken op een verhoogde productiviteit door standaardisatie en efficiëntere *build & self-service deploy*. Onze softwareontwikkelaars kunnen sneller aan de slag. Het gaat ook sneller om iemand nieuw in te werken. Die kan eigenlijk meteen beginnen met coderen", licht hij toe.

De zogenaamde CI/CD-cyclus *Continuous Integration & Continuous Delivery*) wordt hiermee volop ondersteund. "Features kennen een duidelijk kortere doorlooptijd van build tot releasable. We kunnen eenvoudiger en meer continuous releasen."

Tegelijk maakt het containerproject een betere veiligheid en monitoring mogelijk. "We kunnen kwetsbaarheden beter opvolgen en in kaart brengen", stelt hij. "Zo konden we een efficiënte DevSecOps-omgeving uitbouwen dankzij een betere samenwerking binnen onze organisatie."

WAR FOR TALENT

Daarnaast is er een recruitmentaspect. Containers zijn nieuwe technologie, en die oogt goed op de IT-arbeidsmarkt. "We voorzien namelijk een modernere applicatie stack, in lijn met marktevoluties", stelt hij. **"Op die manier is het aantrekkelijker voor IT'ers om bij VDAB te komen werken. Zo'n nieuwe architectuur helpt om deze profielen aan boord te krijgen."**
In die zin acteert VDAB in dezelfde context als alle Vlaamse organisaties waar zij ten dienste van staat. Ten eerste zijn heel wat IT-profielen (van analist-programmeur tot IT-manager) knelpuntberoepen. Ten tweede speelt ook voor IT-profielen de demografische situatie. "Voor alle 100 werknemers die de arbeidsmarkt in Vlaanderen verlaten, komen er slechts 81 bij."

Steven Van Peteghem, directeur I&T operations bij VDAB: "We mikken op een verhoogde productiviteit door standaardisatie en efficiëntere build & self-service deploy."

SHOW ME THE MONEY

Vandaag zijn zowat twee derde van de applicaties en componenten naar containers overgeheveld. Goed voor 160 applicaties en 240 componenten. Het gaat dan voornamelijk om Java- en Oracle-technologie. "Uiteindelijk plannen we 70 à 80 procent over te hevelen naar containers. Eind dit jaar willen we het project afsluiten, al evolueren de applicaties voortdurend en komen er nieuwe functionaliteiten bij."

Helft minder virtuele machines
en 10 miljoen euro besparing.

Het gaat om meer dan een halvering van het aantal virtuele machines: van 4.500 naar 2.229 virtuele machines. Goed voor een kostenbesparing van 10 miljoen op vijf jaar tijd, meer bepaald een lagere TCO door besparingen in licenties en infrastructuur. "**Het voordeel is dan nog iets minder groot dan de 12 miljoen die we eerst hadden ingecalculeerd,** onder meer omdat we de investeringen in tooling wat hadden onderschat", erkent Van Peteghem.

IT'S ALL ABOUT AGILITY

Tegelijk ziet hij 10 procent efficiëntieverbetering. "En een groot pluspunt is dat *onze time to market* gevoelig verbetert. We kunnen bijvoorbeeld sneller data uitwisselen. Dit nieuwe containerplatform helpt dan om zaken sneller opgezet te krijgen en onze wendbaarheid te verhogen", legt hij uit.

Zelf spreekt Van Peteghem over een verbetering in time to market tot bijvoorbeeld 40 procent in *mean-time-to-restore*. Maar hoe stelt zich dat in de praktijk? Hij geeft een voorbeeld. "Wij moesten vanuit de overheid het project 'Mijn weg naar werk', waarbij burgers zicht krijgen op het volledige traject naar het werk, zo snel mogelijk uitvoeren.

Het project integreert onder andere databronnen van VDAB en lokale besturen. Met deze opstelling gaat dat een stuk sneller en efficiënter." Al is het natuurlijk een investering in middelen én mensen. "In totaal is er voor meer dan 2.000 uur aan opleiding, zelfstudie en coaching voorzien."

EN DE LEVERANCIER?

Voor het project ging VDAB in zee met Red Hat, intussen onderdeel van IBM, en zijn OpenShift Container Platform. "RHACS (wat staat voor *Red Hat Advanced Cluster Security for Kubernetes architecture*) is een integraal onderdeel van ons containerplatform", klinkt het bij VDAB.

"Toen we enkele jaren geleden de oefening deden was de keuze voor OpenShift op zogenaamde *bare metal* (waarbij bepaalde software, meestal een besturingssysteem direct op de hardware wordt geïnstalleerd) minder evident. Het argument dat meespeelde, was dat van de vlakke netwerkorganisatie voor medewerkers, samen met de expertise die vandaag rond OpenShift bestaat."

Wat vandaag meer dan toen van tel blijkt, is de onzekerheid rond de nieuwe tarifering van VMware door de nieuwe eigenaar Broadcom. Zelf erkent Van Peteghem ook dat de onderhandelingen met de leverancier hierover best moeizaam kunnen verlopen, een geluid dat al meermaals is opgevangen bij klanten.

EN DE CLOUD DAN?

VDAB gebruikt van het Smals-datacenter in een *on-premise* omgeving. De datagevoeligheid speelt hierbij een rol. De infrastructuur wordt er beheerd door Cegeka.

Ook al is het containerproject een enabler voor mogelijke *move-to-cloud*, vandaag is dat nog niet aan de orde, en zeker de publieke cloud niet. Vandaag beschikt VDAB naar eigen zeggen over een multicloud containerplatform dat op dezelfde leest is gestoeld als in de publieke cloud.

"Al zijn er vanuit Red Hat volop samenwerkingen met de publieke cloudspelers", benadrukt Stef Schampaert, country managing director bij Red Hat Belgium & Luxemburg. Hij verwijst naar initiatieven als *Red Hat OpenShift on AWS* en naar *Red Hat OpenShift on Azure*. "Maar wij zijn met Red Hat geen clouddienstverlener. No way dat alles vandaag naar de publieke cloud kan gaan, zeker niet voor publieke instellingen", stelt Schampaert. "Niet alles gaat cloud native. Meer zelfs: wij merken **dat velen snel terugkomen van het cloud first-verhaal. Kosten spelen hierbij een rol, net als dataprotectie."**

Of zoals Stef Schampaert het treffend samenvat: "Zeg dus niet cloud first, maar cloud applicable. Het open en hybride verhaal speelt veel meer. En VDAB is daar een mooi voorbeeld van."

Stef Schampaert, country managing director bij Red Hat Belgium & Luxemburg: "Niet alles gaat cloud native."

TOCH NOG VIRTUELE SERVERS, EN WEL HIERVOOR

Niet alles kan zomaar naar containers worden overgeheveld. Steven Van Peteghem van VDAB haalt enkele use cases aan waar virtuele servers gewenst en eigenlijk zelfs aangewezen of noodzakelijk zijn:

- **Oracle databases servers.** "Bij de start van het project waren grote, monolithische databases minder geschikt om te draaien op een containerplatform", stelt Van Peteghem.
- **Legacy applicaties.** "Die hebben een ingrijpende re-architecture nodig om te passen op een containerplatform. Of ze draaien op gedateerde software die niet ondersteund wordt op het containerplatform."
- **Buy (commercial-of-the-shelf) applicaties.** "Die zijn technisch vaak niet compatibel met een containerplatform."

WE NEED TO TALK ABOUT DIVERSITY

DIVERSITEIT IN IT: DE HARDE CIJFERS

Man, wit en jong: dat is vaak het clichébeeld van het IT-profiel. Cijfers wijzen uit dat er qua diversiteit in IT nog werk aan de winkel is. Met name in ons land.

17,2%
VAN DE IT'ERS IS EEN VROUW
In Europa bedraagt het gemiddelde aantal vrouwelijke IT-specialisten nauwelijks 17,9 procent, volgens cijfers van Eurostat. Voor België is het beeld nog somberder: slechts 17,2 procent van de IT'ers is een vrouw. Dit gebrek aan diversiteit heeft niet alleen invloed op de bedrijfscultuur, maar kan ook gevolgen hebben voor de innovatie en duurzame groei van de sector.

5%
VAN DE MANAGEMENTROLLEN IN TECH WORDT INGEVULD DOOR EEN VROUW
Het gebrek aan vrouwelijke vertegenwoordiging strekt zich niet alleen uit tot technische functies, maar ook tot leidinggevende posities. Uit Brits onderzoek van PwC blijkt dat slechts een minderheid van de IT-managementrollen in België door vrouwen wordt ingevuld, wat toch wijst op een glazen plafond dat moeilijk te doorbreken lijkt.

0,84%
HET AANDEEL STEM-AFGESTUDEERDEN ONDER VROUWEN
Het aandeel STEM-afgestudeerden in België (een hoger onderwijsdiploma in wetenschap, technologie, engineering of wiskunde) onder vrouwen steeg volgens de laatste cijfers van 7,5 naar 8,4 op elke duizend vrouwelijke 20- tot 29-jarigen. Dat blijkt uit cijfers van de Federale Overheidsdienst Werkgelegenheid, Arbeid en Sociaal Overleg. Daarmee blijven vrouwen duidelijk een minderheid in de STEM-studies, want het aandeel onder mannen is 22,8‰. En ook tegenover de rest van de EU is er nog werk aan de winkel, want gemiddeld zijn er 13,9‰ vrouwelijke STEM-gediplomeerden in de EU.

5%
VAN DE IT-SPECIALISTEN HEEFT EEN NIET-EUROPESE ACHTERGROND
In België heeft slechts 5% van de IT-specialisten een niet-Europese achtergrond, wat relatief laag is tegenover andere sectoren. Ter vergelijking: van de totale Belgische beroepsbevolking bestaat ongeveer 15% uit werknemers van buitenlandse afkomst, waarvan 8,1% afkomstig is van buiten de Europese Unie. Dit betekent dat het aandeel niet-Europese IT'ers significant lager ligt dan het nationale gemiddelde voor niet-EU-werknemers. Dit blijkt uit cijfers van de Hoge Raad voor de Werkgelegenheid.

39 JAAR
Uit cijfers van Agoria blijkt dat de gemiddelde leeftijd van tech- en IT-professionals en digitale profielen in België rond de 39 jaar ligt, met een lage instroom van 50-plussers.

15%
ERVAART LEEFTIJDSDISCRIMINATIE
Volgens een studie van de FOD Werkgelegenheid ervaart 15% van de werknemers ouder dan 45 jaar discriminatie tijdens sollicitatieprocedures. Dit geldt ook voor de technologische sector. In de IT-sector, waar technologie snel evolueert, wordt vaak de voorkeur gegeven aan jongere profielen.

GEMIDDELDE LEEFTIJD BIJ TECHBEDRIJVEN

- **Facebook (Meta):** 28 jaar
- **Google (Alphabet):** 30 jaar
- **Apple:** 31 jaar
- **Amazon:** 31 jaar
- **Microsoft:** 33 jaar
- **IBM:** 38 jaar
- **Oracle:** 39 jaar
- **HP:** 39 jaar
- **Dell:** 35 jaar
- **Cisco:** 34 jaar
- **Intel:** 34 jaar
- **Accenture:** 33 jaar

Deze gegevens tonen dat bedrijven zoals IBM en Oracle oudere medewerkers hebben, terwijl jongere bedrijven zoals Meta een lager gemiddelde hebben. Deze cijfers van StatInvestor dateren wel al van enkele jaren geleden, maar los daarvan geven ze aan dat de gemiddelde leeftijden bij techgiganten niet erg hoog liggen.

MEER VROUWEN IN TECHNOLOGIE?

Volgens tech-topvrouw Tanuja Randery zijn drie zaken van tel om meer vrouwelijk talent in techjobs en -bedrijven te krijgen.

Randery is een van de voorvechters rond diversiteit, ook in haar bedrijf. Ze bekleedt een topjob bij AWS, waar ze VP & Managing Director voor de EMEA-regio is (Europe, Middle East & Africa). Ze is van Indiase afkomst en groeide op in Groot-Brittannië. Ze bekleedde al heel wat leidinggevende functies bij bedrijven zoals Schneider Electric, Colt Technology Services en McKinsey.

"Ik denk dat er nog een weg te gaan is voor de industrie", stelt Tanuja Randery, die zelf een vrouwennetwerk heeft opgezet. "We brengen vrouwen samen om over sommige van deze onderwerpen te praten", stelt ze. "Ik denk dat drie dingen echt belangrijk zijn. Een daarvan is dat we, vanuit het perspectief van de technologie-industrie, **meisjes moeten blijven aanmoedigen om enthousiast te zijn over technologie. En dat betekent echt vroeg beginnen**: op scholen. En niet wachten tot later", stelt Randery. Ze verwijst naar **het eigen Get It-initiatief** waarbij Amazon-werknemers als inclusieambassadeurs met die scholen samenwerken. Dus zij worden de rolmodellen. "Ze komen opdagen en de meisjes zien ons. Ik denk dat het een probleem wordt als we niet jong beginnen."

Maar dit is niet genoeg. Ten tweede moet men vrouwen **aanmoedigen om in tech aan het werk te gaan**. "Dan denk ik aan omscholingsprogramma's, die vrouwen in de digitale economie krijgen door training en opleiding. Ik denk dat dat echt belangrijk is."

Ten slotte is er de loopbaan zelf. "Uiteraard moeten we altijd de structuur voor gelijke beloning tussen mannen en vrouwen goed krijgen", vertelt ze. "Ik denk echt dat het superbelangrijk is binnen onze organisaties dat we niet alleen vrouwen aanwerven in het bedrijf, maar dat we **ze ook helpen hun carrière uit te breiden**."

Randery verwijst naar mentorprogramma's die binnen AWS lopen om vrouwelijke werknemers te helpen. "Van anderen leren, mentorschap krijgen, promotie maken: ,dat soort dingen. Bedrijven moeten ook een inclusieve omgeving hebben waar mensen zich op hun gemak voelen."

Tanuja Randery:
"Er is nog een weg te gaan."

EEN IT-JOB VOOR DE MEEST KWETSBAREN

BECODE IS VEEL MEER DAN CODE

“Als ik vandaag zou beginnen, zou ik misschien voor een andere naam kiezen”, vertelt Béatrice de Mahieu. BeCode is nu immers (veel) meer dan alleen coderen of programmeren. In zijn zesjarig bestaan heeft BeCode al een kleine vierduizend IT’ers opgeleid in een van de vier campussen: Brussel, Gent, Luik en Charleroi.

“Onze oorspronkelijke missie staat nog altijd voorop”, zo opent Béatrice de Mahieu, CEO van BeCode, het interview. “**We zorgen voor de re-krutering, opleiding en begeleiding van de meer kwetsbare profielen** en helpen hen naar werk in de tech-sector”, vertelt ze. BeCodes aanpak is alsmaar meer divers, ook voor de kandidaten. “Dat gaat van het trainen van soft skills van de kandi-daten tot introducties bij mogelijke

werkgevers en administratieve opvolging. Het hele jaar door leveren we kandidaten af."

Wat oorspronkelijk begon met softwaredevelopment is intussen veel meer. "Webontwikkeling doen we bijvoorbeeld nog altijd via een opleidingstraject van zeven maanden. Maar die trajecten zijn er intussen ook voor data & AI, wat nu zelfs onze opleiding met de grootste vraag is. Ook een traject cybersecurity bieden we aan, zowel cybersecurity voor zogenaamde *red* en *blue teams*."

STEEDS MEER OP MAAT

Béatrice de Mahieu benadrukt de opmars van de opleidingen op maat. "Die worden belangrijker. Zo organiseren we binnenkort een opleiding Microsoft Dynamics voor medewerkers van bpost aangevuld met enkele werkzoekenden. Het traject, dat onder meer Power Apps en Dynamics ERP omvat, telt een vijftal certificates. Een ander voorbeeld is ons traject over Cisco-technologie voor een tiental bedrijven die telkens enkele medewerkers afvaardigen", stelt ze.

"We kijken waar bedrijven talent nodig hebben. De opleidingen zijn generiek, maar soms op maat, zoals onlangs een traject *Internet of Things* bij webdevelopment dat we voor Orange Belgium organiseerden", stelt ze. "We schatten vooral ook in. Als verschillende organisaties dezelfde noden hebben, kunnen we dus een opleidingstraject bundelen."

WIE ZIJN DE KANDIDATEN?

BeCode mikt traditioneel op kandidaten die enige afstand hebben tot de arbeidsmarkt. Maar Béatrice De Mahieu geeft wel aan dat iedereen welkom is. "Diversiteit op zich is geen voorwaarde." **90 procent van de kandidaten bestaat wel uit wat je kunt beschouwen als kwetsbare of minder traditionele IT-profielen zoals vrouwelijke kandidaten, vluchtelingen of langdurig zieken.** "Een diploma is niet nodig, maar het kan. Zoals die kandidaat die geschiedenis had gestudeerd en zich omschoolde tot IT'er", haalt ze aan. Naast diversiteit staat ook integratie voorop. "We willen complementair zijn, bijvoorbeeld met het bestaande talent in een organisatie en integreren via bijvoorbeeld stages."

Béatrice de Mahieu, CEO van BeCode: "Je moet vandaag de logica van codering nog altijd inzien. Ook in het tijdperk van ChatGPT."

Voor de instroom van kandidaten hanteert BeCode diverse kanalen. “Dat gebeurt via onze eigen communicatie, bijvoorbeeld via onze sociale media en infosessies. We werken met overheidsorganisaties zoals VDAB, FOREM, Bruxelles Formation en Actiris. Dat doen we ook met sociale organisaties in de diverse steden. We hanteren vaak een andere aanpak per regio”, vertelt ze.

“Code is overal aanwezig, maar het coderen zelf minder.”

Het is dus absoluut geen *one size fits all*-aanpak. “**In Vlaanderen ligt de focus vaak meer op upskilling.** Zo merken we dat voor onze campus in Gent de kandidaten in verhouding wat ouder zijn en gemiddeld ook iets meer diploma’s hebben en vaker uit het buitenland komen, waardoor we soms ook Nederlandse lessen aanbieden. In Luik gaat het dan in verhouding vaker om jonge Belgen zonder diploma.”

WIE ZIJN DE PARTNERS?

Vandaag werkt BeCode voor een veertigtal organisaties, en dit op basis van een jaarlijks contract. “Accenture en Orange waren de eerste partners. **Accenture rekruteerde via ons al 40 BeCoders en bood al 70 stageplaatsen aan.”** Intussen gaat het breed, met namen als NRB, Cegeka, Belfius, Proximus of Ethias.

Het zijn overigens niet allemaal grote organisaties. “Ook reclamebureaus zijn partner, dan gaat het eerder om skills in data of rond Salesforce”, stelt ze. De aangesloten bedrijven betalen jaarlijks een bedrag dat varieert door onder meer de mate van dienstverlening en interactie. Béatrice de Mahieu geeft een jaarlijkse prijsvork van 6.000 tot 15.000 euro die partners betalen voor een samenwerking met BeCode, dat aanpassingen in het curriculum, *use-cases, inspiration talks*, company visits en rekrutering mogelijk maakt.

Dat bedrag is dus erg afhankelijk van die dienstverlening.

IS DE NAAM BECODE NOG VAN TEL?

Eigenlijk steeds minder omdat er in verhouding minder kandidaten zijn voor ontwikkeling dan voor data en AI. “Als ik vandaag zou beginnen, zou ik misschien voor een andere naam kiezen”, zegt ze lachend.

“Code is vandaag overal aanwezig, maar coderen minder. Kijk bijvoorbeeld naar de trend van low code of no code”, stelt ze.

AI blijft het inzicht cruciaal, beklemtoont ze. “Je moet vandaag de logica van codering nog altijd inzien. Ook in het tijdperk van ChatGPT. Je moet misschien zelf niet alle code genereren, maar wel begrijpen wat er gebeurt.”

VIJF BECODE-STARTERS BIJ BELFIUS

Bij Belfius startten deze zomer vijf BeCode-trainees. Na een stage kregen ze er een contract aangeboden. We vroegen Belfius hoe ze dat aanpakken, om welke profielen het gaat en welk inclusiebeleid de bank hanteert.

Bij de aangeworven stagiairs gaat het om vier Java-programmeurs. Zij komen in de *technology bank* teams terecht na een succesvolle stageperiode. En dan is er nog één data & AI-specialist. Die zal bij de bank werken in een van de teams binnen de zogenaamde *business data & digital perimeter*. “De stages duurden een drietal maanden waarin we dus een evaluatie konden doen”, vertelt Gert Vanhaecht, IT-directeur bij Belfius.

SPEERPUNTEN IT RECRUITMENT

Waarom kiest Belfius ook voor deze rekruteringsvorm? “We hanteren een strategie om kritische functies bij voorkeur te laten invullen door interne medewerkers”, stelt Gert Vanhaecht. “We gaan uiteraard prioritair op zoek naar kwalitatieve interne medewerkers. Tegelijk testen we hier met BeCode succesvol een nieuw aanwervingskanaal uit voor resources die we niet altijd eenvoudig vinden.”

Bij dat laatste ligt ook de verklaring voor het gebruik van de diensten van BeCode. “We ontmoetten op deze manier mensen die een prima basisopleiding kregen en duidelijk potentieel hebben om verder te evolueren én die uitermate gemotiveerd en geëngageerd bleken”, stelt hij. “Bovendien bieden we zo kansen aan mensen die op een punt in hun levensloop of carrière kwamen, waar een professioneel gerichte heroriëntatie de juiste stap was.”

Gert Vanhaecht, IT-directeur bij Belfius: “Diverse teams leiden tot de meest optimale resultaten voor onze klanten.”

INCLUSIEBELEID ALS WIN-WIN

Het initiatief past, zo benadrukt Vanhaecht, binnen het inclusiebeleid en het Belfius DNA. “Zo past deze samenwerking in het rijtje waar bijvoorbeeld Passwerk ook in thuis hoort en waarbij we mensen met een divers profiel en andere background kansen willen geven.” Passwerk is een organisatie die consultants aanbiedt, die allemaal een autismespectrumprofiel hebben. Er werken momenteel een tiental consultants van Passwerk bij Belfius.

RESULTATEN

Vanhaecht ziet de inclusieaanpak echt als een win-win. “We weten uit ervaring dat het werken met teams van mensen met een diverse herkomst, interesse of eigenheid tot de meest optimale resultaten leidt voor onze klanten”, vertelt hij.

“We bieden kansen aan mensen die op een punt kwamen waar een professionele heroriëntatie de juiste stap was.”

Bovendien sluit het inclusieve rekruteringsbeleid, volgens hem, ook aan bij de missie van de bank. “Mensen die via BeCode of andere partners een atypisch opleidingstraject volgden, helpen we zo ook mee een toekomst te geven. Zo vullen we ook de maatschappelijke rol in die we voor ons weggelegd zien. Dat zit ook in onze zogenaamde purpose. Die luidt: *Meaningful & Inspiring for the Belgian Society. Together.”*

GROOTSTE DISCRIMINATIE IN IT: LEEFTIJD

Als er op de arbeidsmarkt één hardnekkige vorm van discriminatie blijft opspelen, dan is het ten aanzien van oudere kandidaten. En dat is zeker ook bij IT-profielen van tel. “Het lijkt er alleen maar hardnekkiger op te worden.”

Uit onderzoek van UGent blijkt dat **oudere jobkandidaten gemiddeld 40 procent minder positieve reacties** kregen op hun sollicitaties dan volstrekt gelijkwaardige kandidaten met een jongere leeftijd. “In tegenstelling tot etnische discriminatie lijkt leeftijdsdiscriminatie ook helemaal niet af te nemen over de tijd”, constateert **Stijn Baert**, professor arbeidseconomie aan UGent.
In de Verenigde Staten speelt leeftijdsdiscriminatie veel minder. Terwijl de leeftijdsverschillen in de Amerikaanse praktijktesten een pak hoger lagen dan bij die in Europa.

Let wel, leeftijd is een vooraanstaande vorm van discriminatie, maar niet de enige. Zo kunnen, ook in Vlaanderen, een arbeidsbeperking of uiterlijke kenmerken ongeveer dezelfde mate van discriminatie veroorzaken. Maar discriminatie op basis van seksuele geaardheid wordt bij ons niet echt vastgesteld, volgens Baert. Ook een bewijs voor genderdiscriminatie bij aanwerving is er niet (uitgesproken). “**Discussies over diversiteit op**

de arbeidsmarkt gaan bijna altijd over achterstelling van allochtonen en vrouwen. Het lijkt op die manier alsof andere groepen geen achterstelling kennen. Of toch veel minder. En dat is fout", stelt Baert, die als academicus regelmatig leeftijdsdiscriminatie aankaart.

TWEE VOORBEELDEN UIT DE IT-WERELD

Ook in IT speelt het aspect leeftijd heel erg. Maar je merkt het pas als je het nader bekijkt. Zo moesten bij de recente ontslagrondes van technologiebedrijven als Cisco en SAP in ons land toch vooral de oudere profielen het gelag betalen. Bij SAP kreeg enkele maanden geleden bijna 10 procent van het lokale personeelsbestand te horen dat ze hun koffers mochten pakken. **Opvallend was dat bij de ontslagen vooral de ietwat oudere medewerkers zaten.** "Zo goed als iedereen ouder dan 55 zou moeten vertrekken", zo klonk het in de berichtgeving.

Ondanks wat je zou verwachten, neemt leeftijdsdiscriminatie niet af bij hogere krapte op de arbeidsmarkt. En ook niet bij IT. Vaak lijken er ook geen gegronde redenen voor, zo blijkt uit een tweede praktijkvoorbeeld. "Ik had in mijn opleiding toegepaste informatica een vijftiger. De student had al een diploma industrieel ingenieur en was een van de beste studenten van zijn jaar", vertelt **Jan Guldentops**, docent informatica aan de AP Hogeschool. "Maar toch had die student het – net omdat hij al iets ouder was – uiterst moeilijk om aan een stageplek te geraken. Ik vond dat frappant", stelt hij. "*Ageism* is ook in de IT-jobmarkt een groot en onderbelicht probleem. En het lijkt alleen maar hardnekkiger te worden."

WAT ZIJN DE REDENEN?

Waarom is leeftijdsdiscriminatie zo hardnekkig, zelfs bij knelpuntberoepen zoals IT'ers? "Dat is eigen aan een combinatie van factoren", stelt **Frank Vander Sijpe**, director HR Trends & Insights bij HR-adviseur Securex. "Enerzijds is een job in IT qua kennis en vaardigheden onderhevig aan de snelle, regelmatige en ingrijpende veranderingen in technologie", stelt hij.

"Ik had in mijn opleiding een vijftiger, een van de beste studenten van zijn jaar. Toch had die student het – net omdat hij al iets ouder was – uiterst moeilijk om aan een stageplek te geraken. Ik vond dat frappant."

JAN GULDENTOPS,
DOCENT INFORMATICA AAN
DE AP HOGESCHOOL

Als IT'er moet je constant bijblijven, en vaak knelt daar het schoentje. "Het is een gekend fenomeen, zeker in België, dat de leerspanning afneemt met de leeftijd. Ervaring staat dus niet altijd garant voor kennisvernieuwing", weet Vander Sijpe. "In feite wordt aangenomen dat jongere werknemers niet alleen goedkoper zijn, maar ook meer up-to-date qua kennis. Ze worden ook al snel geacht om flexibeler te zijn in het aanleren van nieuwe zaken."

Toch zijn dit niet de enige redenen. **Eigenlijk werkt het huidige arbeidsmarktsysteem die leeftijdsdiscriminatie ook een beetje in de hand**. "Ons loonsysteem is gebaseerd is op anciënniteit-gebonden en baremieke loonsverhogingen. Verhogingen die vaak te weinig of zelfs helemaal niet gerelateerd zijn aan prestaties of werkattitude. Eigenlijk is dit een typisch Belgische situatie, die in andere landen veel minder speelt."

Pittig detail: wettelijk mag je als werkgever niet discrimineren bij aanwervingen. "Maar je zou als werkgever wel heel dom zijn als je je kandidaat meegeeft dat zijn hogere leeftijd de reden is voor een eventueel afwijzen", stelt hij. "Daarom is discriminatie per definitie iets dat stiekem gebeurt."

"Men neemt snel aan dat jongere werknemers niet alleen goedkoper zijn, maar ook meer up-to-date qua kennis", zegt Frank Vander Sijpe, director HR Trends & Insights bij HR-adviseur Securex.

DRIE VRAGEN OVER JOU EN JOUW IT-CARRIÈRE

1

HOE WAPENEN TEGEN LEEFTIJDSDISCRIMINATIE?

Hoe kun je je als IT'er wapenen tegen leeftijdsdiscriminatie? "Er zijn verschillende opties", stelt Frank Vander Sijpe van Securex. "Je kunt ofwel 'breed' gaan ofwel je specialiseren in één bepaald domein, zoals SAP. Maar aan elke keuze zijn risico's verbonden. Als je breed gaat, dan kun je je plan trekken in een aantal expertises, maar moet je passen eens werkelijk diepgang gevraagd wordt." Maar ook die andere aanpak houdt risico's in. "Als je diep gaat, dan ben je een expert. Maar dat ben je maar tot zolang een bepaalde expertise relevant blijft en niet vervangen wordt door een andere."

Los hiervan heeft Vander Sijpe nog twee suggesties. "Een **vaak gehoorde of toegepaste aanpak is die van het zogenaamde T-profiel**. Zo'n *T-shaped professional* beschikt over specialistische kennis op een bepaald terrein, maar heeft zich ook breed ontwikkeld." Ten tweede wijst hij op de competenties van de toekomst. "Leerbereidheid en flexibiliteit. Jongere medewerkers hebben de reputatie om op dat vlak minder dogmatisch te zijn. De competenties van de toekomst lijken in het nadeel te spelen van de oudere IT'ers, al mag je ze uiteraard niet allemaal over dezelfde kam scheren."

2

WELK STATUUT: VAST OF EXTERN?

Ben je als IT'er niet beter af in een zelfstandig statuut dan als (vaste) werknemer met een oplopende (en voor de werkgever dure) anciënniteit? Dat hangt ervan af, ook in functie van de conjunctuur. "Vanuit een standpunt van een zelfstandig IT'er zijn het vandaag gouden tijden. Hun statuut geeft de mogelijkheid om te kiezen welke opdrachten ze wel of niet willen aannemen. Je bevindt je niet in een hiërarchisch ondergeschikte relatie, zoals dat wel het geval is voor een vaste medewerker", stelt hij.

Als freelancer zit je meer in de relatie van een partnership. Zolang het loopt en je goed in de markt zit, gaat dat goed. Heel goed zelfs. "**Vaak krijgen de freelancers zelfs dan de mooie IT-projecten, terwijl de vaste IT-werknemers het moeten stellen met het minder interessante cyclische IT-gedoe**. Werkgevers hebben bij schaarste de neiging om goede freelancers toch zo veel mogelijk aan zich te binden", stelt hij. "Als freelancer komt het er dan ook op aan om volgende projecten en prospecten in de pipeline te hebben. Je moet je namelijk kunnen vermarkten."

3

EN DE WERKGEVER ZELF?

Veel werkgevers en opdrachtgevers huiveren er nog voor om strategische IT-projecten aan freelancers over te laten. Terecht, vindt Vander Sijpe. "Voor zo'n opdrachtgever stelt de vraag zich of de kennis en vaardigheden al dan niet noodzakelijk zijn voor de core business", oppert hij. "Ook werken met freelancers biedt niet altijd de garantie van stabiliteit qua prijzen, en kan een bedrijf bovendien in een soort afhankelijkheidspositie duwen, met alle gevolgen van dien."

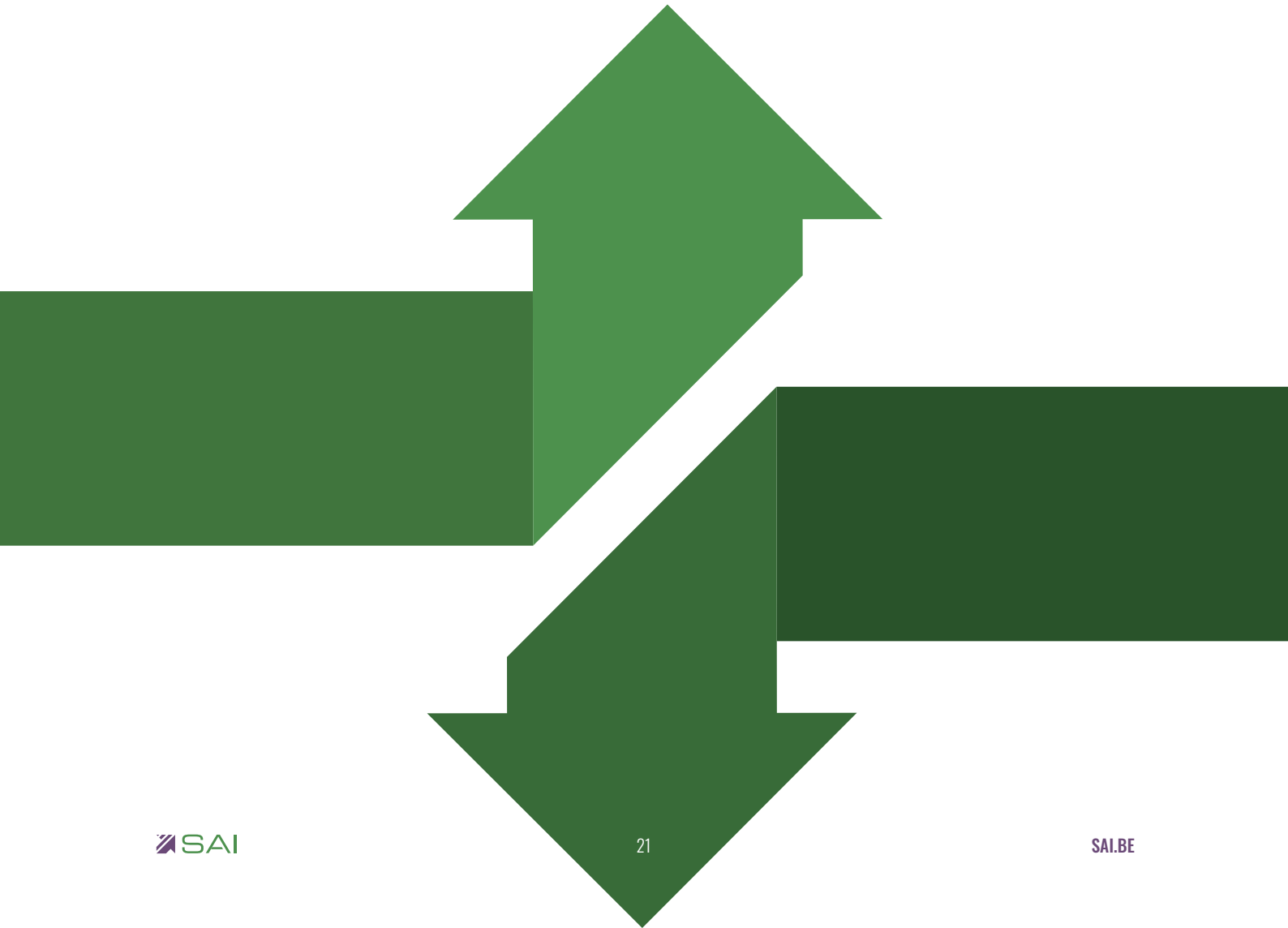
EN DAN NU HET GOEDE NIEUWS: DE OPENSTAANDE IT-JOBS IN DE KOMENDE VIJF JAAR

Uit recent onderzoek van Steunpunt Werk, in verband met de aanwervingsbehoeftes in Vlaanderen naar sector en beroep, blijkt dat er meer IT-profielen nodig zullen zijn. In de periode 2024-2029 zullen volgens hun projecties 6.400 jobopeningen nodig zijn in IT, tegenover 2.250 in de periode 2014-2023. Misschien ironisch in deze context: vooral de grote vervangingsgraad van oudere profielen (55+) is hiervoor de grootste oorzaak.

FUTUREPROOF

WELKE IT-SKILLS ZITTEN IN DE LIFT? EN WELKE ZIJN OP HUN RETOUR?

Als IT'er moet je goed in de markt te liggen, of je nu vast of als externe kracht in dienst bent. SAI Update brengt je een overzicht van de vaardigheden die het vandaag goed doen. Ook de dalers lijsten we op. Hiervoor namen we contact op met twee rekruteringsbureaus die gespecialiseerd zijn in IT-profielen.



Recruiters komen met verschillende werkgevers in contact. Ze hebben dan ook een goed beeld van wat er op de arbeidsmarkt gevraagd wordt. “Uit recent onderzoek, dat we deden voor onze Salarisgids, komen er enkele interessante trends naar boven rond skills”, vertelt **Jeroen Diels**, regional managing director bij Robert Half.

ALGEMENE TRENDS

Jeroen Diels bekijkt het eerst breed. “We merken dat kennis van programmeertalen, degelijke knowhow van AI en machine learning om bijvoorbeeld processen te vereenvoudigen of te automatiseren, business intelligence en rapportering veelgevraagde technische vaardigheden zijn”, stelt Diels. Hij maakt bewust een onderscheid tussen **enerzijds de hard skills en anderzijds de soft skills**. “Communicatieve vaardigheden, probleemoplossend vermogen en aanpassingsvermogen zijn gegeerde soft skills.”

Niet geheel verrassend blijkt dat artificiële intelligentie ook hot is op de arbeidsmarkt. “Zo zijn kennis van AI en machine learning eveneens moeilijk te vinden skills”, vult collega **Wout Van de Wijer**, manager bij Robert Half, aan. “Als we kijken naar vakgebieden met grote vraag naar IT-specialisten, dan komen vooral AI, IT service support, IT architecture en IT security naar voren. **Het is op zich weinig verrassend dat alles rond AI sterk gevraagd wordt**, gezien de opmars van de toepassingen.”

Qua domeinen of skills die op hun retour zijn, klinkt het algemene antwoord genuanceerd. “De IT-sector is traditioneel een sector die doorheen de jaren talloze veranderingen ondergaat: nieuwe technologieën die zich in een razendsnel tempo ontwikkelen en daarmee gekende toepassingen op losse schroeven zet en minder en minder zal ondersteunen. Denk hierbij aan technologieën zoals AS/400, Lotus Notes, Cobol en Progress”, haalt Diels aan. “**Qua specifieke functies zien we voorlopig niet meteen een gigantisch verschil of afbouw van een bepaald type rol** in het bijzonder. De laatste jaren zien we zelfs het tegenovergestelde: een uitbreiding en diversifiëring van het veld en de functies door technologische ontwikkelingen.”

“Artificiële intelligentie is ook hot op de arbeidsmarkt”, stelt Jeroen Diels, regional managing director bij Robert Half.

STIJGERS EN DALERS PER VAKDOMEIN

1/ ICT BUSINESS (FUNCTIONELE & BUSINESS ANALYSE, PROJECT MANAGEMENT, ...)

Welke skills zitten in de lift?

"We zien de laatste jaren de vraag stijgen naar BI en data analytics posities", vertelt **Céline Luypaert**, recruitment manager ICT Business bij Amandis. "Bedrijven willen een beter inzicht krijgen in hun data om (strategische) beslissingen te onderbouwen en gaan daarom meer op zoek naar profielen met ervaring in data-analyse en BI-tools, zoals Power BI)", vertelt Luypaert. "En inderdaad: ook meer en meer bedrijven zetten in op AI."

Daarnaast liggen, volgens haar, skills voor *change management* goed in de markt. "Dit komt door bedrijven die digitale transformaties (moeten) doorvoeren en die verandering moeten kunnen begeleiden. Denk aan functies zoals de transformation consultant of de change manager."

Tot slot, voegt ze eraan toe, zijn veel bedrijven overgeschakeld van de klassieke *waterfall*-methodologie naar agile. "Hierdoor zijn er extra functies bijgekomen zoals scrum masters of product owners. Profielen die ervaring hebben met agile en scrummethodologieën zijn dan ook gewild."

Welke zijn op hun retour?

Traditioneel waterfall-projectmanagement heeft zijn beste tijd gehad. "Vroeger werkten bedrijven vaak via de waterfall-methodologieën, maar de afgelopen jaren zagen we een duidelijke verschuiving naar de agile methodologie", merkt Céline Luypaert op. "Dit is te wijten is aan de rigiditeit van waterfall en aan de toenemende behoefte aan wendbaarheid in bedrijven. Agile is veel flexibeler. Projectmanagers die zich alleen richten op waterfall zonder kennis van agile, worden hierdoor minder vaak gezocht."

"Veel bedrijven zijn overgeschakeld van de klassieke waterfall-methodologie naar agile", vertelt Céline Luypaert, recruitment manager ICT Business bij Amandis.

Brent Vermeulen, recruitment manager software & data bij Amandis: "DevOps-methodologieën zoals Jenkins, Docker en Kubernetes zijn cruciaal om softwareontwikkeling te versnellen."

2/ SOFTWARE DEVELOPMENT & DATA

Welke skills zitten in de lift?

Binnen softwareontwikkeling kun je verschillende vaardigheden opsplitsen, stelt **Brent Vermeulen**, recruitment manager software & data bij Amandis.

- Backend development: "De 'standaardtechnologieën' zoals Java, .NET, PHP en Python blijven zeer populair in deze markt. Anderzijds merk je een sterke toename van Node.js en Golang (Go)", stelt Vermeulen.
- Frontend development: "React, Angular en Vue blijven de populairste frameworks binnen frontend."
- Mobile development: "Er is meer nood aan mobile developers dan vroeger. Technologieën zoals Flutter, React Native en Kotlin zijn populairder geworden."

Bovendien migreren bedrijven volgens Brent Vermeulen massaal naar de cloud, waardoor ontwikkelaars ervaring moeten hebben met AWS, Azure of Google Cloud. "Voorts zijn DevOps-methodologieën zoals Jenkins, Docker en Kubernetes cruciaal om de softwareontwikkeling te versnellen en de productkwaliteit te verbeteren."

Welke zijn op hun retour?

"Er is veel minder vraag naar verouderde programmeertalen zoals Cobol, Delphi, of Pascal", haalt Brent vermeulen aan. "Vroeger waren WinForms en WPF ook zeer populair om desktopapplicaties te bouwen, maar tegenwoordig is de focus meer verschoven naar web- en mobiele applicaties."

3/ INFRASTRUCTURE, SUPPORT & SECURITY

Welke skills zitten in de lift?

Vandaag zijn vaardigheden in cloud computing, zoals ervaring met AWS, Microsoft Azure en Google Cloud, zeer gewild door de populariteit van cloudtechnologieën, is de vaststelling van **Vincent Meert**, senior recruitment consultant infra, support & security bij Amandis. “Cloudproviders nemen veel van de traditionele infrastructuurtaken over, waardoor cloudkennis een integraal onderdeel is geworden van moderne infra IT-vacatures.” Ook cybersecurity is intussen een belangrijke skill. “In dit domein is er een sterke vraag naar expertise in netwerkbeveiliging, threat detection en compliance, wegens de toename van cyberaanvallen.”

Welke zijn op hun retour?

Met de verschuiving naar cloudoplossingen is de vraag naar beheerders voor fysieke servers en lokale datacenters afgenomen. “Dit betekent echter niet dat deze rol verdwijnt, maar eerder dat ze evolueert”, nuanceert Vincent Meert. “De traditionele systeembeheerder transformeert naar een hybride system engineer. Deze professionals moeten nu kennis hebben van cloudplatforms en automatiseringstools.”

Vincent Meert, senior recruitment consultant infra, support & security bij Amandis: “De traditionele systeembeheerder transformeert naar een hybride system engineer.”

WELKE CERTIFICATEN LIGGEN ERG GOED IN DE MARKT?

Op basis van de input van Robert Half en Amandis krijgen we deze opsomming:

- Information Technology Infrastructure Library (ITIL)
- Agile of scrum certificaten zoals Certified ScrumMaster (CSM)
- PRINCE2 (traditioneel, maar blijft waardevol)
- DevOps zoals Docker Certified Associate, Certified Kubernetes Administrator (CKO)
- Softwarearchitectuur zoals TOGAF 9 Certified
- Cloud computing-certificaten zoals AWS, Azure en GCP
- Cybersecurity zoals CISSP, Certified Ehtical Hacker (CEH) en CompTIA Security+

DE ENIGE WACHTWOORDREGEL DIE OVEREIND BLIJFT

Lange tijd was het credo: wachtwoorden zijn als onderkleding. Je moet ze regelmatig aanpassen. Maar de opmars van multifactorauthenticatie doet het beleid voor IT-beheerders veranderen. “Eigenlijk is er vandaag nog maar één regel van tel bij wachtwoordbeleid: laat gebruikers nooit zelf hun wachtwoorden kiezen.”

1

WACHTWOORDEN BLIJVEN VAN TEL

Terwijl organisaties jarenlang sterke wachtwoorden hebben afgedwongen, lijkt dat in het tijdperk van multifactorauthenticatie (waarbij bijvoorbeeld een extra code of bevestiging op de smartphone van de eindgebruiker nodig is bij het loginproces) minder van tel.

Onterecht, vindt Reinaert Van de Cruys, ethical hacker en medeoprichter van Fox & Fish Cyberdefense. Hij komt bij veel organisaties over de vloer en neemt er hun securitybeleid onder de loep. "**In tal van gemeentes zijn ze nog volop bezig met de uitrol van multifactorauthenticatie (MFA). In veel private bedrijven overigens ook**", vertelt hij tijdens het webinar van SAI.BE. "Zij hebben dan bijvoorbeeld Remote Desktop Protocol (RDP) of webmail op internet staan zonder MFA, alleen met een wachtwoord beveiligd."

Bovendien zullen er nog lang en vaak toepassingen blijven bestaan, waar geen multifactorauthenticatie bij gebeurt. Hierdoor blijft de nood aan een slimme wachtwoordpolicy bestaan.

2

DE KLASSIEKE REGELS HELPEN NIET (MEER)

Een veilig wachtwoord verandert regelmatig, is minstens twaalf tekens lang en bevat letters, cijfers en symbolen. "In 80 procent van de bedrijven waar ik kom, hebben ze nog dergelijke wachtwoordregels. Maar we trappen als IT'ers in de valkuil dat als je over een hamer beschikt, ieder probleem eruitziet als een nagel", oppert Van de Cruys. "Wij hebben de voorbije dertig jaar als IT'ers onze gebruikers gedwongen om veilige wachtwoorden te kiezen, en dat op manieren die eenvoudig te checken zijn", stelt hij.

Maar eigenlijk is het ook belangrijk dat een wachtwoord geen informatie bevat over jou als gebruiker. Dat er geen info over iemand instaat – van hobby, huisdier, kinderen of postcode tot partner – is als IT-beheerder al moeilijker te controleren bij jouw gebruikers. "Met als gevolg dat iedereen zijn wachtwoord *Laura 2014!* blijkt te zijn. En *Laura 2014!* is geen veilig wachtwoord." Er spelen nog andere zaken mee. Als jouw wachtwoord bijvoorbeeld ook de naam van jouw kinderen bevat, en het maakt deel uit van een datalek, dan is dat nog extra informatie die wordt gelekt.

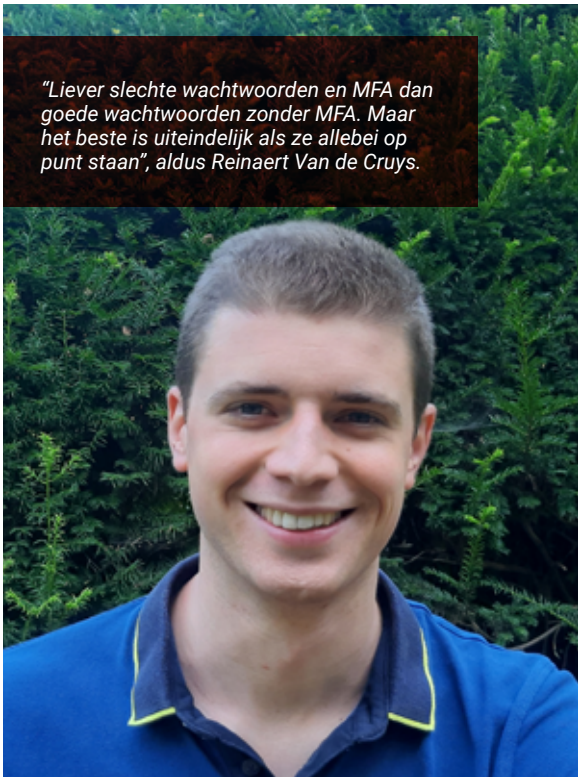
Bill Burr, manager bij het National Institute of Standards and Technology – eigenlijk de vader van de wachtwoordregels – gaf dat onlangs zelf ook toe. Hij gaf de aanzet om mensen te dwingen om veilige wachtwoordregels te kiezen. "Na twintig jaar zijn we er eindelijk in geslaagd om mensen wachtwoorden te laten kiezen die moeilijk te onthouden zijn, maar gemakkelijk te raden", zo moest Burr toegeven. "**Bij veel IT-organisaties erkennen de IT-beheerders zelf dat de helft van de medewerkers slechte wachtwoorden hebben. Terwijl ze er niet eens een structurele oplossing voor zoeken**", voegt Van de Cruys toe.

3

WACHTWOORDEN VERANDEREN, WERKT VAAK AVERECHTS

En regelmatig veranderen? "Dat kan een goed idee zijn", stelt hij. Als wachtwoorden lekken op het darkweb, dan is het een goed om tijdig een ander wachtwoord aan te maken. "Al kom ik zelf in veel IT-organisaties waar zelfs de systeembeheerder zijn wachtwoord aanpast naar van *Laura 2014!* naar *Laura 2015!*. En dit op basis van een regel die hij zelf heeft ingevoerd."

Bovendien lijkt het regelmatig veranderen van wachtwoorden vaak een pervers effect met zich mee te brengen. "Ik stel vast dat het mensen ertoe brengt om hun wachtwoorden minder serieus te nemen. Sommigen



"Liever slechte wachtwoorden en MFA dan goede wachtwoorden zonder MFA. Maar het beste is uiteindelijk als ze allebei op punt staan", aldus Reinaert Van de Cruys.

gaan ervan uit dat ze hun wachtwoord met een collega kunnen delen, omdat het toch verandert over drie maanden. Een blijvend wachtwoord is daarentegen 'hun grote geheim'", aldus Van de Cruys. Hij ziet nog een andere reden om wachtwoorden niet of niet vaak te veranderen. "Het leven van je medewerkers en de IT-systeembeheerders wordt er gemakkelijker door. Het levert minder supporttickets op en in de praktijk ben je veiliger."

4

DE BESTE REGEL: WACHTWOORDEN NIET ZELF KIEZEN

De beste, en eigenlijk meest vooraanstaande, regel voor wachtwoorden en wachtwoordbeleid vanuit IT is volgens hem: kies de wachtwoorden niet zelf. "En dat geldt ook voor de IT-beheerders. Kom nooit in de verleiding om zelf snel een wachtwoord te verzinnen." Zelf opteert Van de Cruys voor generators zoals *Last Pass* of *Password Creator*. Wachtwoordzinnen vindt hij oké, maar wel wat omslachtig. "Maar je kunt ook de eerste letters van het woord uit die zin gebruiken."

"Laura 2014! is geen veilig wachtwoord."

5

MULTIFACTORAUTHENTICATIE LOST NIET ALLES OP

In het tijdperk van multifactorauthenticatie zijn wachtwoorden minder van tel. "Multifactorauthenticatie is een goede oplossing als het goed wordt aangepakt", stelt hij. Er zijn ook verschillen en recente hacks tonen dat ook aan. Een bekend fenomeen is *multi factor fatigue*, waarbij hij verwijst naar de hacks bij Cisco en Uber, hoewel die MFA gebruikten. "Hackers beschikten over het wachtwoord van de medewerker en stuurden hem tientallen verzoeken om binnen te raken. Die kwamen op de smartphone bij die medewerker terecht met de vraag: was jij het? "Na tientallen keren op 'neen' te hebben gedrukt, dacht de medewerker dat het om een technisch probleem ging en klikt die op 'ja' om ervan af te zijn. Met alle gevolgen van dien."

Een alternatief voor multifactorauthenticatie is om telkens een bepaalde code over te typen. "Maar hier dook weleens een zogenaamde *man in the middle* attack op." Dan wordt bijvoorbeeld een medewerker opgebeld en aan de hand van *social engineering* begeleid aan de telefoon. De medewerkers gaven de code door aan de telefoon. **"IT'ers zijn technisch onderlegd en kun je moeilijk social engineeren, maar tegelijk staan ze met veel meer mensen in contact**. Een gevoelige vraag valt bij hen soms minder op. En de impact is vaak ook veel groter."

Een ander voorbeeld was het lek bij Last Pas, waar hackers zijn binnengeraakt via een vulnerability op een onvoldoende beveiligde Apple-thuiscomputer van een medewerker. "Daar was het een kwestie van keyloggers en een toestel was geïnfecteerd. Als dat het geval is, smelt de bescherming van MFA als sneeuw voor de zon."

6

DE WINNENDE COMBINATIE

De conclusie is dat MFA is een belangrijke beveiligingslaag is. "Liever slechte wachtwoorden en MFA dan goede wachtwoorden zonder", zo erkent hij. Maar MFA alleen is niet voldoende. Het beste is uiteindelijk als ze allebei op punt staan. En liefst nog meer. Een combinatie van beveiligingsmethodes, training, awareness en up-to-date technologieën is noodzakelijk om aanvallen te voorkomen. Noem het gerust de winnende combinatie.

» *Herbekijk **hier** als SAI.BE-lid gratis het SAI.BE-webinar met Reinaert Van de Cruys.*

HERGEBRUIK VAN SOFTWARECOMPONENTEN: AGILE EN KOSTENEFFICIËNT TEGELIJK

HET WARM WATER IN SOFTWAREONTWIKKELING

Warm water moet je niet opnieuw uitvinden. Dat geldt ook voor software. Je kunt gerust verder ontwikkelen op basis van herbruikbare componenten. Dat is het idee achter het ReUse-initiatief.

Vaak is softwareontwikkeling maatwerk, terwijl dat eigenlijk niet nodig is. Al is het maar omdat de behoeften voor bepaalde toepassingen vaak vergelijkbaar zijn. Zeker ook bij openbare instellingen is dat het geval.

Daarom lanceerde Smals, de gemeenschappelijke ICT-organisatie van Belgische overheidsdiensten in de sociale zekerheid, samen met partners het ReUse-initiatief dat hergebruik van softwarecomponenten promoot. “De partners ondersteunen het ReUse-initiatief, hebben bijgedragen aan de catalogus of bieden soortgelijke initiatieven aan. Partners zijn natuurlijk ook gebruikers”, vertelt Florence Bruyère van Smals.

DAAR IS DE CATALOGUS

De online catalogus bevat intussen meer dan honderd herbruikbare componenten. Het gaat voornamelijk om stand-alone API’s, maar je vindt er ook producten en libraries in terug. Het staat allemaal op **hun website,** die recent werd vernieuwd. “De ReUse-catalogus wordt voortdurend bijgewerkt: sommige

componenten worden verwijderd, andere worden toegevoegd op basis van vereisten, projecten en IT-ontwikkelingen."

Standaardcomponenten of generieke services kunnen bijvoorbeeld gebruikt worden voor het beheren van bestanden, het inloggen, het controleren van de gegevenskwaliteit van informatie, het veilig verzenden ervan, het opvragen van gegevens uit een authentieke bron, formulieren enzovoort. Deze zijn allemaal ontwikkeld met hergebruik in het achterhoofd. "Over het algemeen zijn de ReUse-gebruikers zeer gevarieerd. Dat gaat van instellingen of bedrijven tot burgers", vertelt Florence Bruyère.

"Softwarehergebruik is vandaag een integraal onderdeel van onze projectmethodologie", zegt Florence Bruyère van Smals.

MILJOENENBESPARING

Hergebruik van software is een integraal onderdeel geworden van de werking van Smals, in het bijzonder bij hun softwareontwikkeling. Aan het begin van een projectcyclus bekijken teams welke functionaliteit gerealiseerd kan worden met bestaande componenten en services, welke nieuwe ontwikkeling generiek genoeg is voor toekomstig hergebruik en welk deel volledig maatwerk blijft.

53%

DE ROI OP HERBRUIKBARE SOFTWARE

Gemiddeld bespaart deze werkwijze meer dan de helft van het projectbudget volgens Smals. "In 2023 was het rendement op de investering (ROI) in hergebruik ongeveer 53 procent. Ondertussen worden er steeds nieuwe diensten en componenten ontwikkeld met hergebruik in gedachten", stelt Bruyère. "**In totaal is ongeveer 65 procent van alle functionaliteit gerealiseerd in het kader van hergebruik**, met een geschatte besparing van meer dan 44 miljoen euro in 2023."

SOFTWARE OP ÉÉN WEEK TIJD

Door te mikken op het hergebruik van software kan een overheidsdienst zoals de RSZ (Rijksdienst voor Sociale Zekerheid) niet alleen besparen, maar ook sneller schakelen. Bij het uitbreken van de covidcrisis zelfs in één week tijd.

"Socialezekerheidsinstellingen zoals de RSZ zijn de grootste bijdragers van herbruikbare componenten", vertelt Florence Bruyère van Smals. Meer dan duizend medewerkers bij de RSZ gebruiken een centrale en gestandaardiseerde dossierbeheertoepassing, een zogenaamde werkomgeving of kortweg wo. De medewerkers vinden er onder meer een lijst van taken die ze kunnen uitvoeren. De taken zijn op hun beurt een onderdeel van digitale processen, die doorstromen tussen meerdere teams, of over instellingen heen. Overzichtsschermen geven een overzicht van de uitgevoerde en op te volgen zaken, individueel en per team. Het is een geheel aan generieke componenten voor het opzetten van dossierbeheer, het koppelen van authentieke bronnen en de orkestratie van processen. De wo maakt deel uit van de ReUse-catalogus met componenten, services en producten. Hiermee vormt het een onderdeel van het programma voor softwarehergebruik binnen de Belgische sociale zekerheid en hun gemeenschappelijke ICT-organisatie.

COVID

Door deze aanpak is een nieuw proces flexibel en snel op te zetten, klinkt het. **Bij het uitbreken van de covidcrisis zelfs in één week tijd, voor het uitstel van RSZ-bijdragen voor meer dan honderdduizend ondernemingen.** Dit is mogelijk door generieke functionaliteiten bij elkaar te brengen en vaker een beroep te doen op configuratieparameters dan op specifieke softwareontwikkeling. "ReUse laat toe om snel en goedkoper nieuwe toepassingen op te zetten. Op basis van de bouwstenen van de wo heeft ook de Rijksdienst voor Arbeidsvoorziening (RVA) ondertussen een dossierbeheertoepassing opgezet", klinkt het bij Smals. Het is, volgens Smals, een primeur dat een centrale softwaretoepassing van dergelijke schaal generiek werd opgezet voor meerdere teams en zelfs voor meerdere instellingen. "Dit maakt het eenvoudig om snel en betaalbaar nieuwe administratieve processen op te zetten, of te hervormen." Het plan was en is dat op termijn de wo zelfs een onderdeel kan worden van een referentiearchitectuur voor overheden, en van transversale processen over meerdere instellingen heen.

NIEUWE SOFTWARE OOK HERBRUIKBAAR

Het initiatief is gericht op de promotie van generieke software, bouwstenen en online diensten die overheden en hun leveranciers kunnen hergebruiken in hun ICT-ontwikkeling. "Bij Smals maakt softwarehergebruik al enkele jaren deel uit van de standaard projectmethodologie", beaamt Florence Bruyère. **Jaarlijks wordt twee derde van de functionaliteit via ReUse gerealiseerd**, met zo'n vijftig procent kostenbesparing als ROI. Voorts werd zes procent van de nieuwe ontwikkeling opnieuw herbruikbaar gemaakt. De wo is dus opgevat als een herbruikbare 'service' voor de RSZ, de RVA en andere diensten. Door het grote aantal betrokken teams, aan de business-kant en bij de ontwikkeling van de wo, zijn afspraken, standaarden en een open architectuur essentieel, volgens de initiatiefnemers. "Ook is een grote rol weggelegd voor de 'producteigenaar', die nauw contact houdt met de teams en gebruikers. Er wordt intensief op een agile manier gewerkt, met wekelijkse technische releases en devops-praktijken."

AI EN CYBERSECURITY MAKEN HET OUDE MEDIUM WEER POPULAIR

DE COMEBACK VAN TAPEOPSLAG

Tapeopslag, een technologie die al decennia bestaat, lijkt zijn comeback te maken in het tijdperk van big data en AI. Waar tapeopslag in het verleden al herhaaldelijk dood werd verklaard, lijkt het vandaag aan populariteit te winnen.

In 2023 werd er maar liefst 512,9 exabyte aan tapeopslag verscheept, een stijging van 3 procent tegenover het voorgaande jaar. HPE, IBM en Quantum – de belangrijkste tapeproducenten – rapporteren voor het derde jaar op rij een stijging in de verkoop. Het gaat hierbij meer bepaald om de LTO-technologie, voornamelijk gebruikt voor gegevensarchivering.

WETGEVING EN AI

Wat zit er achter deze opmerkelijke heropleving van een opslagmedium dat al jaren als 'dood en begraven' wordt omschreven? Het feit dat organisaties veel data voor lange tijd (moeten) bewaren, en liefst offline. "Het meest voorkomende gebruik van tapeopslag is het opslaan van archiefgegevens die niet meer regelmatig worden geraadpleegd", aldus Jason Stagnitto van Cloudwards.

Vandaag heeft zowat alles in IT een AI-link, en dat is zelfs met tape niet anders. **De hernieuwde populariteit van tapeopslag kun je ook toeschrijven aan de explosieve groei in datavolumes, aangedreven door AI-workloads.** Dit alles creëert gigantische hoeveelheden ongestructureerde data die ergens veilig en kostenefficiënt opgeslagen moeten worden. Tape blijkt hierin een geschikte op-

lossing voor langetermijnopslag van 'koude data', oftewel gegevens die niet regelmatig geraadpleegd worden. Dit type data kan jarenlang op tape worden bewaard, tegen aanzienlijk lagere kosten dan bijvoorbeeld harde schijven (HDD's) of solid-state drives (SSD's).

RANSOMWARE

Hoewel moderne opslagmedia zoals SSD's en de cloud veel snellere toegang tot data bieden, blijft tapeopslag aantrekkelijk wegens de lage kosten en de relatief hoge duurzaamheid. Tape kan gegevens tot wel dertig jaar veilig bewaren, zonder dat het medium veroudert of significant aan betrouwbaarheid verliest. Dit maakt het ideaal voor back-ups en archivering, vooral in sectoren waar databehoud van cruciaal belang is, zoals de financiële sector, overheidsinstellingen en de gezondheidszorg.

Een ander voordeel van tape is dat het bestand is tegen ransomwareaanvallen, aangezien tapes gemakkelijk offline bewaard kunnen worden. Dit biedt een extra beveiligingslaag tegen cybercriminaliteit, wat vandaag een belangrijk voordeel is.

Voor bedrijven die duurzaamheid hoog in het vaandel dragen, **biedt tape een groenere oplossing voor databeheer**. Uit **gegevens van de Enterprise Storage Group** blijkt dat tape libraries een 6,5 keer minder CO_2 uitstoten ten opzichte van een disk, simpelweg door 500 TB aan gegevens van schijf naar tape te verplaatsen.

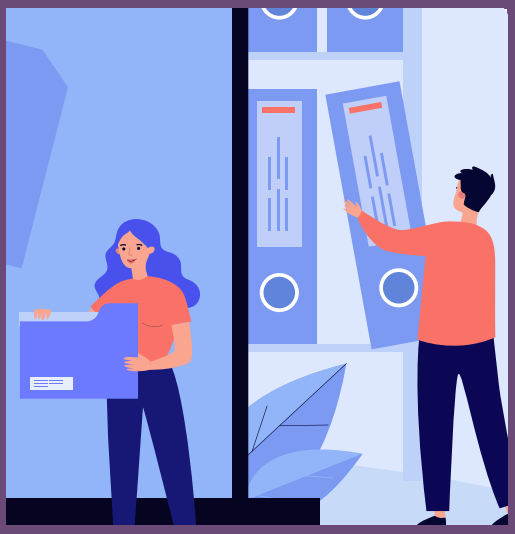
NADELEN

Tapeopslag kent, ondanks de recente heropleving, ook duidelijke nadelen voor bedrijven. **Een van de grootste beperkingen is de lage toegankelijkheid**. Tape is ontworpen voor 'koude' data die zelden opgevraagd worden. Het kan traag en complex zijn om gegevens terug te halen.

In tegenstelling tot harde schijven of cloudopslag, waar data direct beschikbaar is, vereist tape fysieke toegang en specifieke hardware om de gegevens te lezen, wat kan leiden tot aanzienlijke downtijd bij herstel of toegang tot kritieke data. "Het belangrijkste verschil tussen een tapedrive en een harde schijf is nu eenmaal dat een tapedrive sequentiële toegangsopslag gebruikt. Data worden in een specifieke, geordende volgorde gerangschikt in plaats van willekeurig, zoals bij een harde schijf", aldus Jason Stagnitto van Cloudwards. Het gevolg is dat de toegang tot gegevens op tapedrives een aanzienlijk langzamer proces is.

De infrastructuur voor tapeopslag vereist ook gespecialiseerde apparatuur, zoals tape drives en robots voor automatische tape libraries, wat bijdraagt aan de **hoge initiële investeringskosten** en de noodzaak voor gespecialiseerd IT-personeel voor onderhoud en beheer. Tape heeft daarnaast geen toegangsmogelijkheden in real time zoals moderne systemen, waardoor het voor sommige use cases, zoals dagelijkse operaties en kritische workloads, minder geschikt is.

Bovendien sluit tape minder goed aan op moderne datamanagementstrategieën, zoals cloudgebaseerde systemen en flexibele, schaalbare opslagoplossingen, waardoor bedrijven met hybride omgevingen extra complexiteit moeten beheren. Tapeopslag kan daarom een lastige keuze zijn voor bedrijven die snelle toegang en hoge beschikbaarheid van gegevens eisen. Al zijn er voldoende use cases die tape vandaag wel (en zelfs nog meer) aantrekkelijk maken.



TAPE VERSUS DISK: DE MARKT

Hoewel de totale hoeveelheden opgeslagen data door HDD's en SSD's nog altijd aanzienlijk groter zijn, blijft de groei in tapeopslag opvallend. In 2023 werd er wereldwijd 152,9 exabyte aan tapeopslagcapaciteit verscheept door de grote tapeproducenten, zoals HPE, IBM en Quantum.

Ter vergelijking: harddriveproducent Seagate alleen al verscheepte 99 exabyte aan schijfopslag in één kwartaal van 2024. Dit wijst erop dat er dit jaar wereldwijd duizenden exabytes (zetabytes) aan HDD- en SSD-opslag worden verscheept. En het benadrukt dat tape nog altijd een nichemarkt bedient, maar wel een markt die blijft groeien. En ook eentje met een zekere legacy. **Schattingen suggereren dat tapeopslag momenteel verantwoordelijk is voor ongeveer 10% tot 20% van de wereldwijde langetermijnopslag.**

VOORDELEN

Zelfs in de wereld van cloud en disk biedt tapeopslag nog altijd enkele voordelen voor zijn gebruikers.

1. **Levensduur:** Tapes zijn robuust en kunnen tientallen jaren meegaan, met een relatief laag risico op beschadiging van gegevens of schijfdefecten, zeker wanneer de technologie in goede omstandigheden wordt bewaard. Wel blijven tapes gevoelig voor slijtage, hitte, vocht en magnetische velden.

2. **Kosten:** De kosten per gigabyte om data met tapeopslag te bewaren zijn laag in vergelijking met andere opslagopties, waardoor het een efficiënte manier is om oudere, gearchiveerde gegevens op te slaan.

3. **Veiligheid:** Tapeopslag is veiliger dan andere soorten opslag voor het beschermen van gegevens tegen cybercriminaliteit, omdat moderne cyberdreigingen geen toegang hebben tot de data op een magnetische tape.

NADELEN

Tape is (zeker) niet voor elke opslagvorm de beste oplossing: er zijn een paar nadelen aan deze aanpak.

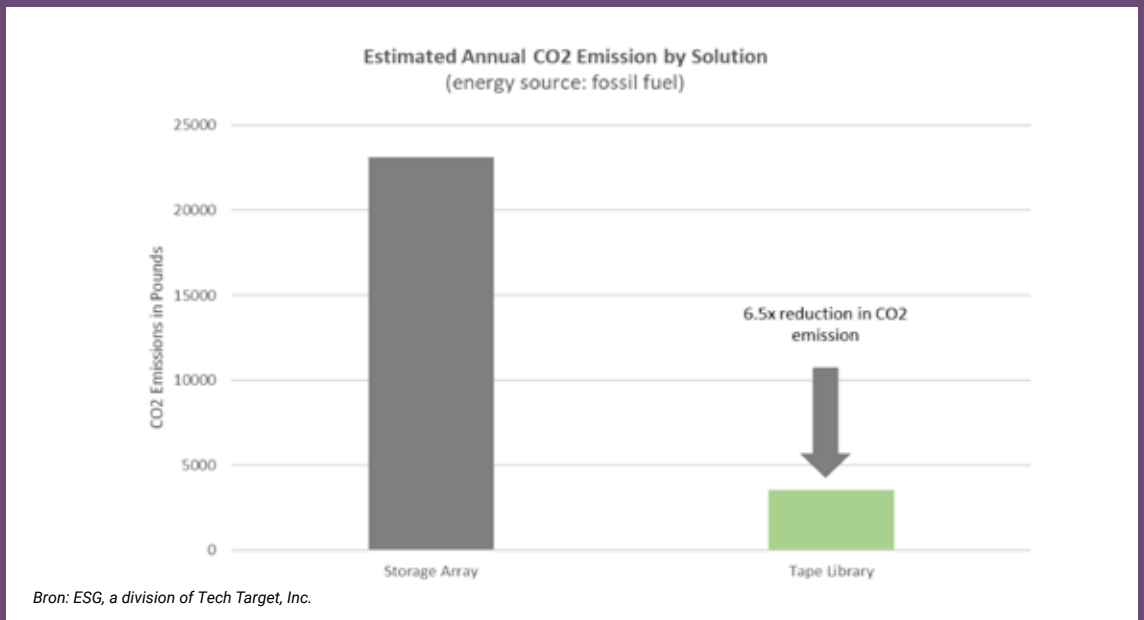
1. **Toegankelijkheid:** Data die snel en regelmatig toegankelijk moeten zijn, kun je beter bewaren met andere back-upoplossingen, zoals een harde schijf of cloudopslag.

2. **Snelheid:** De lees- of kopieersnelheid van tapedrives is relatief laag, vooral in vergelijking met de snelheden van harde schijven of cloudservices.

3. **Fysieke kosten:** Hoewel de kosten per gigabyte laag zijn, kunnen de kosten van de infrastructuur en de tapes zelf behoorlijk hoog oplopen.

HOE GROEN IS TAPE?

Uit recent onderzoek van de Enterprise Storage Group blijkt dat tape libraries (het apparaat dat tapes beheert en automatiseert) 6,5 keer minder CO_2 uitstoten dan een storage array, een opslagsysteem dat meerdere disks (HDD's of SSD's) samenvoegt in één apparaat.



Bron: ESG, a division of Tech Target, Inc.

WAT KUNNEN WE LEREN UIT DE HACK VAN DE EEUW?

In het internationale true crime-boek *De hack van de eeuw* zet misdaadjournalist Diamant Salihu uiteen hoe een internationale politiesamenwerking gecodeerde berichtendiensten wist te hacken, waardoor ze maandenlang en in real time mee konden lezen met de communicatie tussen drugsbazen, koeriers en tienerhuurmoordenaars.

Het boek van Diamant Salihu leest als een thriller, met veel personages en geen al te beste afloop. True crime dus. Wij halen uit de 317 pagina's enkele elementen, en dan vooral wat ons als IT'ers kan interesseren.

1/ NIET ÉÉN MAAR MEERDERE HACKS

De hack van de eeuw zijn er eigenlijk verschillende. In onze contreien deed vooral de kraak van de Sky ECC-berichtendienst veel stof opwaaien. Het zette een heuse internationale klopjacht tegen drugscriminelen in gang. Dankzij het inkijken van de berichtendienst konden politiemensen namelijk aan heel veel inside-informatie geraken. Maar eigenlijk werden er dus meerdere berichtendiensten gekraakt.

In het boek van Diamant Salihu gaat het vooral om de EncroChat die gehackt kon worden dankzij een Frans-Nederlandse politiesamenwerking. Het duurde 73 dagen vooraleer EncroChat zelf opmerkte dat de berichten van hun dienst op straat lagen.

Daarna ging het om de Anom-berichtendienst die werd ontmanteld (met behulp van de FBI). Dit was een berichtendienst die ironisch genoeg door de criminele onderwereld zelf op het getouw was gezet. Daarna stapten veel criminelen over naar **Sky ECC, dat veiligheid beloofde en een miljoenenbedrag uitloofde aan degene die de dienst kon kraken. Maar ook die dienst ging dus door de knieën.**

2. NET ALS HET ENIGMA VAN WOII

"Het is net zoiets als Enigma in de Tweede Wereldoorlog", zo vertelt een politieman in het boek over de kraak van EncroChat. Het breken van de Enigmacodes van de toenmalige elektromechanische codeermachines bleek toen een goudmijn aan informatie op te leveren. De informatie, verkregen door ontcijfering van de geheime Duitse berichten, speelde een uiterst belangrijke rol in het verloop van de Tweede Wereldoorlog.

In het boek van Salihu hebben politiemensen het over een *once in een lifetime*-gebeurtenis toen ze EncroChat kraakten. "Beter dan dit wordt het niet", zo klonk het. De berichtenstroom was internationaal en de inhoud sloeg de politiediensten met verstomming. Ook in Zweden, waar een groot deel van het boek zich afspeelt en waar ze heel erg met drugscriminaliteit te maken hebben (het is zelfs een bekend exportproduct voor het land).

Overal schrikt de politie naar aanleiding van de kraak hoeveel drugs het land binnenkomen. Want drugs is big business. De Nederlandse – en intussen opgepakte – drugscrimineel Ridouan Taghi beheerste op een bepaald moment een derde van de Europese cocaïnesmokkel. Hij is (of was) goed voor een vermogen van maar liefst een miljard euro.

3/ ANDROID-TELEFOON MET EEN PANIEKKNOP

Volgens Franse overheidsdocumenten had de versleutelingsdienst EncroChat op een bepaald moment 60.000 gebruikers, hoofdzakelijk in Europa. Dat waren niet allemaal drugscriminelen, al maakten ze er wel gretig gebruik van.

De toestellen hadden eigenlijk een spiegelfunctie. Ze zagen eruit als een ordinaire Android-telefoon, en zo werkten ze ook. Maar het

andere besturingssysteem bevatte EncroChat. Zodra je een wachtwoord had ingetypt opende de cryptowereld zich, een wereld waar je veilige berichten kon sturen.

Het smartphonesysteem heeft ook een veiligheidsfunctie, mocht de gebruiker bijvoorbeeld door een politie-inval verrast worden. Dan kun je **een paniekcode invullen om de volledige inhoud van je telefoon te wissen.**

4/ DATACENTER IN ROUBAIX

Toen het cybercrimeteam van de Franse gendarmerie de Encro-telefoons nauwkeurig geanalyseerd had, deden ze een interessante ontdekking, die later cruciaal zou zijn voor het verdere verloop van het verhaal. Al het dataverkeer bleek namelijk een datacenter in het Franse Roubaix te passeren. Dat gaf de Franse aanklagers de mogelijkheid om een opsporingsonderzoek tegen rechtspersonen te beginnen, in dit geval tegen telecombedrijf EncroChat. Andere landen, en hun politiediensten, zullen hier later op aansluiten.

5/ NOG JAREN AAN DATA-ANALYSE VOOR POLITIE

Veel politiediensten, en zeker die in Zweden, gebruiken voor de analyse van de vele gekraakte berichten het platform *Acus Table*, een data-programma dat grote hoeveelheden data kan verwerken. Alleen al Sky ECC bestaat uit 35 miljoen unieke berichten, en dat gaat dan nog uitsluitend om de mededelingen die een link hebben met Zweden. De politiediensten hebben nog enkele jaren te gaan om al het Sky ECC-materiaal door te nemen, als dat al lukt.

6/ DE IRONIE VAN DUBAI

In het hele drugsverhaal speelt Dubai een rol van betekenis. Veel criminelen trekken er naartoe om uit de klauwen van de Europese politiediensten te blijven. Zo kunnen ze de dans ontspringen. Maar ironisch genoeg zijn ze daar in de Arabische Emiraten behoorlijk streng als het om misdaad en criminaliteit gaat. De gevluchte criminelen durven er geen wapens bij zich te hebben of iemand anders iets aan te doen, want dat kan daar voor hen namelijk de doodstraf betekenen.

"Het duurde 73 dagen voor ze bij EncroChat zelf opmerkten dat hun dienst gekraakt was."

Intussen is Dubai op internationaal vlak iets meegaander geworden. Af en toe wordt er een crimineel opgepakt en het land uitgezet, zoals Ridouan Taghi, die in Dubai bij de lurven werd gevat. Tijdens de politie-inval zat Taghi met zijn vriendin in de zetel en keek tv. Het ingrijpen van de politie ging razendsnel en zonder bloedverlies. Het enige verlies was de rode wijn op het kamertapijt, wijn die Taghi op dat moment in zijn schuilplaats aan het degusteren was.

7/ SCHURKENSTAAT TURKIJE

Dan kun je vooral Turkije aanzien als *gangster paradise*. Veel criminelen vluchten naar Turkije om gerechtelijke vervolging in Europa te ontlopen. En als ze dan toch voor de rechter komen in Turkije, kopen ze hun proces vaak af of minimaliseren ze de bewijslast.

Turkije is een natie waar je het Turks staatsburgerschap simpelweg kunt aankopen, door bijvoorbeeld voor minstens 400.000 euro vastgoed te kopen in het land. Degenen die goud of buitenlands kapitaal invoeren, hoeven zich niet te verantwoorden waar het geld vandaan komt. "Dat maakt Turkije uiterst aantrekkelijk voor het witwassen van drugsgeld'", schrijft Diamant Salihu in zijn boek.

8/ DE DOOS VAN PANDORA IS OPENGEGAAN

Een van de hoofdstukken uit het boek heeft als titel *De doos van Pandora*, naar een verhaal uit de Griekse mythologie. Dat gaat over het huwelijkscadeautje dat Epimetheus en Pandora cadeau kregen, en niet mocht geopend worden. Daarin zaten alle ongelukken van de wereld opgesloten.

Eigenlijk was dit met de hack deels ook het geval. Omdat bepaalde gangsterbenden ontmaskerd werden en achter de tralies verdwenen, ontstond er turbulentie in de drugsscène. Er verschenen nieuwe groeperingen op het toneel, de ene al driester (en jonger) dan de andere. Helemaal onderaan de ladder staan de 'kindsoldaten' en moord-commando's klaar om opdrachten uit te voeren.

Drugsdelinquenten geven zelf aan dat steeds meer onbesuisde minderjarigen worden ingezet, omdat de strengere wapenwetgeving ertoe leidt dat oudere criminelen minder risico's willen lopen. Hierdoor is het geweld alleen maar geëscaleerd. Dat is eigenlijk nog de meest sombere conclusie uit deze hack van de eeuw.

De hack van de eeuw *- Hoe de politie live mee kon lezen met de grootste drugscriminelen van Europa, Diamant Salihu, Xander Uitgevers, ISBN: 9789401622646, juli 2024, 318 pagina's.*

OPVALLENDE QUOTES

“Het vermogen om goed met AI te werken en er je voordeel mee te doen, is nu belangrijker dan het begrijpen van Excel of het internet.”

Bill Gates, in een interview met de Amerikaanse techsite CNet, naar aanleiding van de lancering van 'What's next, the future with Bill Gates', zijn docuserie op Netflix.

“Best number to reach me is € 1000 daily rate.”

Freelance IT'er Pieter Debaere op LinkedIn in een reactie op een bericht van een headhunter met een ‘exciting Data Analyst position for a huge energy client’ en de vraag: ‘What’s the best number to reach you in to discuss further?’

“Dit neigt naar maffiapraktijken.”

SAI.BE-voorzitter Marc Vael in aflevering 38 van de SAI.NT-podcast over de verkoops- en handelsaanpak van softwaregigant Adobe, dat onduidelijk communiceert over de kosten van zijn (moeilijk opzegbaar maandelijks) abonnement. Adobe werd er ook voor aangeklaagd door het Amerikaanse ministerie van Justitie.

“Wanneer alle belanghebbenden er niet in slagen om vanaf het begin in te stemmen, zijn zerotrustinitiatieven gedoemd om te mislukken.”

Ismael Valenzuela, senior instructor bij Sans Institute, op Computable. “Effectieve beveiliging wordt aangestuurd door processen en technologie, maar vooral door mensen”, benadrukt hij.

“Sinds Pornhub in Louisiana de leeftijd controleert, is het surfverkeer op de site met 80 procent gezakt.”

Een woordvoerder van Aylo, het bedrijf boven adultsite Pornhub, naar aanleiding van de discussie over verplichte leeftijdscontrole op bepaalde websites. Volgens hem “trekken surfers naar donkere hoeken van het web waar je leeftijd niet wordt gecheckt”.

Linus Torvalds

“C is, in the end, a very simple language. And because it’s simple, it’s also very easy to make mistakes.”

Linux-oprichter Linus Torvalds op The Register over de discussie om de programmeertaal Rust te gebruiken voor de Linuxkernel.

DATA VISUALIZATION

FINDING THE FINGERPRINT OF DATA

Data visualization is the one of the modern work horses of data science. Why is this so important? The simple reason is that, if the dimensionality of the data becomes too large, then the data might be difficult to visualize in a traditional way. Well, what is meant by this?

Suppose that you are working at a telecom operator and you are tasked with the responsibility of finding out how the telecom service can be more profitable. Recently, a new competitor entered into the market. However, the telecom operator still wants to be competitive and wants to find out how they can keep clients.

To make this task easier the telecom operator provides information about every client they recently served or are serving. The information the operator keeps from each customer are features. These can be the number of call failures, subscription length, seconds of use, frequency of use, frequency of SMS, distinct called numbers, if it is a pay-to-go or contract, amount of data usage, age of the customer and ultimately if they recently changed provider or are still active.

This complete collection of client-information is a dataset. One instance/datapoint of this dataset is hence a **client** and every client **has some features**. The amount of features is called the **'dimension' of the data**.

A **datapoint or instance** P corresponding to a certain observed instance with **measured features** $p_1, \cdots, p_n$ can be presented in the following way:

$$P = [p_1 \ p_2 \ p_3 \ p_4 \cdots \ p_{n-2} \ p_{n-1} \ p_n]$$

Note that this datapoint has dimension n. This is just one datapoint, so this means that the **complete dataset is a matrix**. For N clients the complete data-matrix looks as follows:

$$\boldsymbol{P} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_N \end{bmatrix} = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} \cdots p_{1,n} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \cdots p_{2,n} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} \cdots p_{3,n} \\ & & \vdots & \\ p_{N,1} & p_{N,2} & p_{N,3} & p_{N,4} \cdots p_{N,n} \end{bmatrix}$$

To get back to the client example, one will try to see if there can be **predicted** based on the other features that a client will possibly leave the deal or will stay. Hence, one needs to **classify** a customer in 'leaving' or 'staying'.

This is a task that can be done by a support vector machine or neural network. A full explanation of these deserves an article on its own. One must remember that a support vector machine (SVM) is a very reliable tool for binary classification, where one typically classifies in class 0 or class 1, or in this case in class 'client left the deal' or 'client stayed in the deal'. A neural network can theoretically do the same thing as a SVM. Both have their pros and cons.

For now, the goal is to **distinguish leaving clients from present clients**. In this way, there can be discovered which things in the service need to be improved to remain competitive.

First, we need to discover if there are serious differences between leaving clients and present clients before looking into improving services. Enter visualisation.

HIGH DIMENSIONALITY

Humans live in a three-dimensional world. This means that we can perfectly visualize lines, surfaces and volumes. However, **anything with higher dimension we cannot really imagine**. How does something look like in 4D? This seems more difficult.

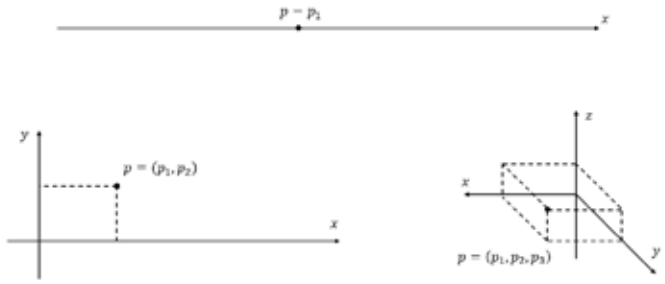


Figure 1: A single number can be represented on a line, two numbers on a surface and three numbers in a volume, but what with four numbers or more?

Hence, if we have a collection of datapoints with a higher dimension than three, **we cannot visualize it in a traditional way.** If the data cannot be visualized in the traditional technique due to the fact there are more than three features, then we have difficulties. It will be much harder to see if there is a pattern that can be exploited or two (or more) classes that can be easily separated.

FINDING A SOLUTION

A first proposal is **projecting the high-dimensional data on a line**. The problem with this method however is that datapoints which are far away from each other can be projected close to each other in the lower-dimension representation. This is not something which is desired because it does not give a reliable interpretation of the data.

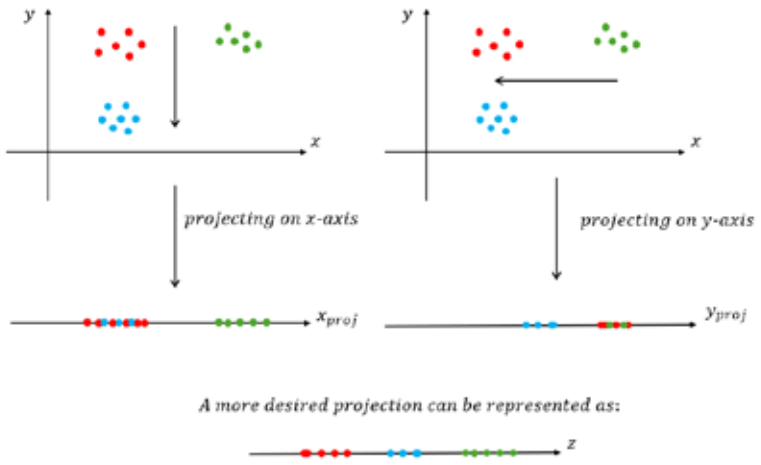


Figure 2: When projecting on the x-axis and the y-axis, it is clear that lower-dimensional representation is retrieved, which does not give a reliable representation of the data. Something more desired is shown below x_{proj} and y_{proj}.

t-SNE (t-distributed stochastic neighbour embedding)

Enter t-SNE**, a technique used in general to visualize high-dimensional datasets**. Let's start with a similar dataset as shown in figure 2. Then after running t-SNE in MATLAB the following is achieved:

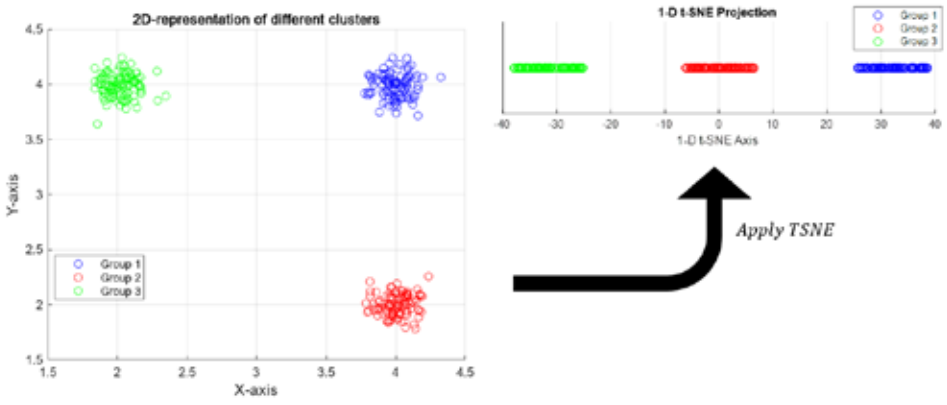


Figure 3: original dataset (left), 1D-dataset after running t-SNE (right).

How is the 1D-representation of the 2D-data achieved? First, the different datapoints are drawn on a 1D-line in random order. Note that since this is determined randomly. Hence, t-SNE will give a different representation for each time one performs it. Then the position of each datapoint is updated so that it is attracted to points close to it in the original dataset and it is repelled from datapoints which are far from the point in the original dataset.

First, there needs to be found out how much each point is similar to another point. This is done using a function g which used the distance between two points p_i and p_j in the given space:

$$s_{ij} = g(\|p_i - p_j\|)$$

What we want here is a value s_{ij} that is between 0 and 1 and represents **how similar points are**. Here 0 means that i and j are unsimilar and 1 means that they are the same point. Hence, something presented below is desirable:

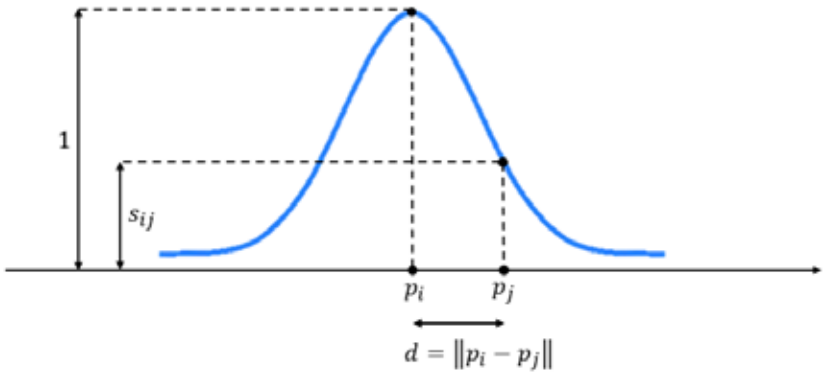


Figure 4: What is desired is that points close to p_i express high similarity. Points further from p_i express smaller similarity.

It is important to note that there are still some things to consider. For example, what if there are clusters which are **less dense**? Then it might be that these are underrepresented in the final representation.

In addition, if this rule is applied then there should hold that $s_{ij} = s_{ji}$. It should make sense that the similarity of point i to j is the same as the similarity of the point j to i.

To counteract the effect of less dense clusters one first calculates the following:

$$s_{ij} = \frac{g(\|p_i - p_j\|)}{\sum_{i \neq k} g(\|p_i - p_k\|)}$$

In this way, clusters with different densities will be taken into account. In this way there holds $\sum_{j=1, i \neq j}^{N} s_{ij} = 1$.

BEHIND THE CURTAINS

When one then looks at the paper of t-SNE, it turns out that the similarity is calculated as:

$$s_{ij} = \frac{\exp\left(-\|p_i - p_j\|^2 / 2\sigma_i^2\right)}{\sum_{i \neq k} \exp(-\|p_i - p_k\|^2 / 2\sigma_k^2)}$$

The function g we use here is a Gaussian. Why do we use this function? Well, it is a **good measure to see if two points are similar,** and it completely satisfies the desired shape presented in figure 4. This function will give points located closely to each other a higher similarity value than points far away in the considered (high-dimensional) space.

Using this, it turns out that the demanded symmetry is not present anymore. To counteract this, one then performs:

$$S_{ij} = \frac{s_{ij} + s_{ji}}{2n}$$

Here n is the amount of datapoints present in the dataset.

After calculating all the similarities between the points one can obtain the following matrix $\underline{S}$.

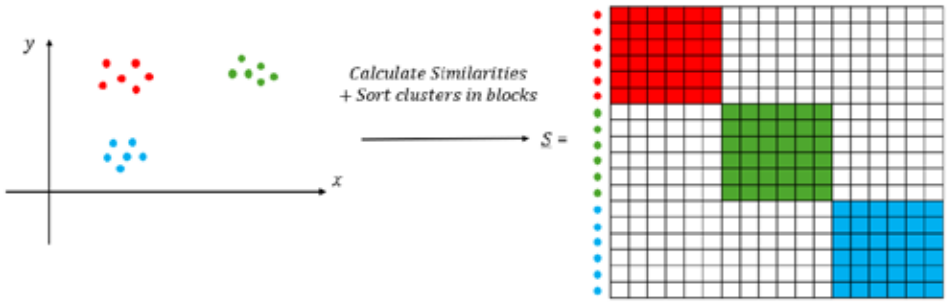


Figure 5: Clusters can be found when ordering the similarities in blocks.

Since the similarities are large between points in the same clusters, they can be ordered in these blocks. Thus, each block represents a cluster. The white values represent values which are smaller than the coloured values. This means that the similarity between different clusters is much smaller than the similarity between points in the same cluster.

Getting the lower-dimensional representation

Now that this step is done. It is time to **map from a larger dimension to a smaller dimension**. As mentioned before, all points are randomly initialized in the lower dimension.

We will randomly distribute the datapoints in a lower-dimension space like a line, surface or volume. Then shift the points along the space until a desired outcome in the lower-dimensional representation is achieved.

First, we start measuring the distance in the lower dimension to construct a similarity measure c_{ij} in the lower dimension between de points q_i and q_j using the similarity function h.

$$c_{ij} = \frac{h(\|q_i - q_j\|)}{\sum_{i \neq k} h(\|q_i - q_k\|)}$$

Since the points are all randomly initialized, the similarity matrix C won't look anything like the nice matrix, so we need to shift each point at the time so that it looks more like the original matrix $\underline{S}$. This is similar like a **puzzle that needs to be solved**. We need to shift each point at the time so that in the end we end up with the same distribution as in the original dimension.

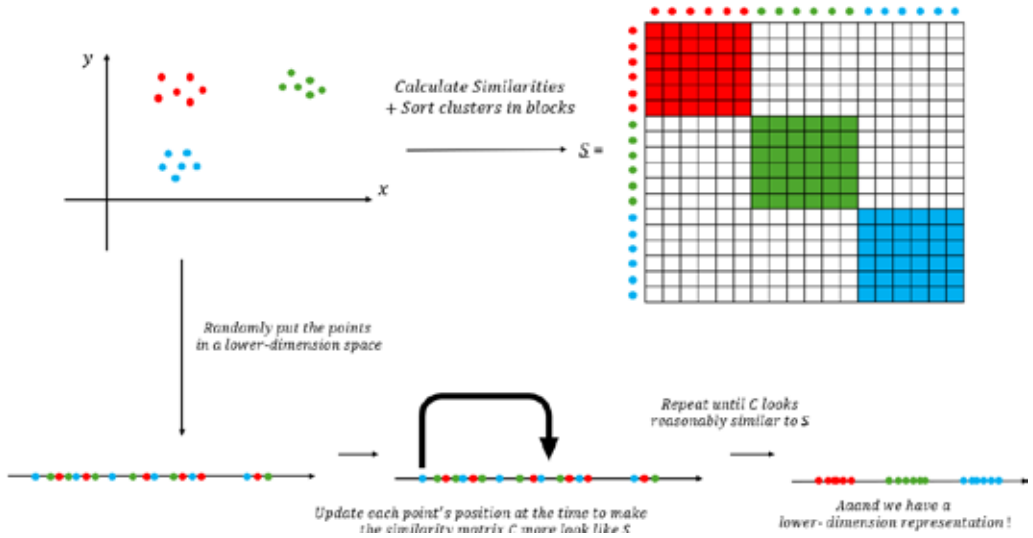


Figure 6: complete overview of the process.

There is one sidenote here. Note that it was left unclear how h could be chosen. One could argue that again the Gaussian function could be chosen. This would give the SNE-algorithm. However, **there is a good reason why we would use a different kind of similarity measure**.

If we would use a Gaussian function, then the low-dimensional representation might be too crowded. The different clusters would all be clumped together, and the visualization would not be desirable.

Why this is the case can be explained through an analogy. Imagine you have a huge room, like a ballroom, where all your friends are hanging out. In this big room, there's plenty of space, so your friends can spread out. Some are chatting in small groups, some are standing alone, and some are far away from each other.

Now, imagine you want to move all your friends into a much smaller room. In this smaller room, you still want to show who was close to who in the big room. But because the room is much smaller, everyone has to stand closer together.

The problem is, when you squish everyone into the small room, **you can't keep the same amount of space between them as they had in the big room**. Even though two friends were far apart in the ballroom, they might end up standing next to each other in the small room because there's not enough space. This makes it hard to show the same arrangement of friends as in the big room, and that's the "**crowding problem**". In a lower-dimensional space there is simply less space.

Hence, we need to make sure that things spread out and to do this we use a distribution with heavier tails namely the student-t distribution.

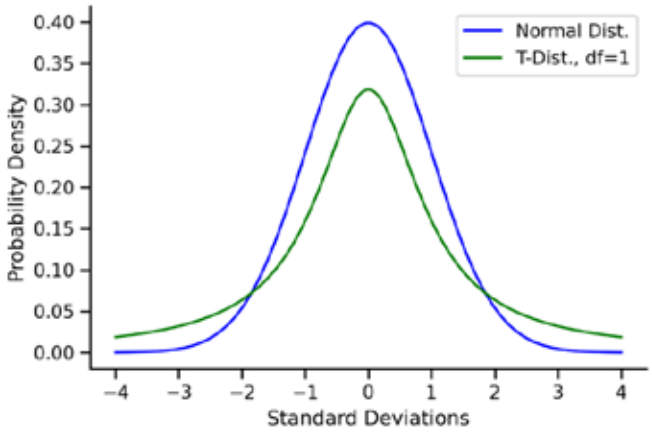


Figure 7: The tails of the student t-distribution are larger than the tails of a Gaussian distribution.

Imagine you have a special rule in the small room: friends who were far apart in the big room still need a bit more space around them, even if everyone has to get closer together. The heavy tails of the Student-t distribution act like this special rule, **making sure that distant friends don't get too crowded together in the small room**. This helps preventing the crowding problem, where everyone would otherwise end up too close to each other, making it hard to see the original relationships.

Hence, in the end the following is done. Note that this similarity function h is similar to the desired shape of a similarity function presented in figure 5.

$$c_{ij} = \frac{\left(1 + \|q_i - q_j\|^2\right)^{-1}}{\sum_{i \neq k}(1 + \|q_i - q_k\|^2)^{-1}}$$

RESULTS OF t-SNE

Now that we went through all the dominant aspects of t-SNE, it is important to see how it does for some cases. Suppose the following dataset:

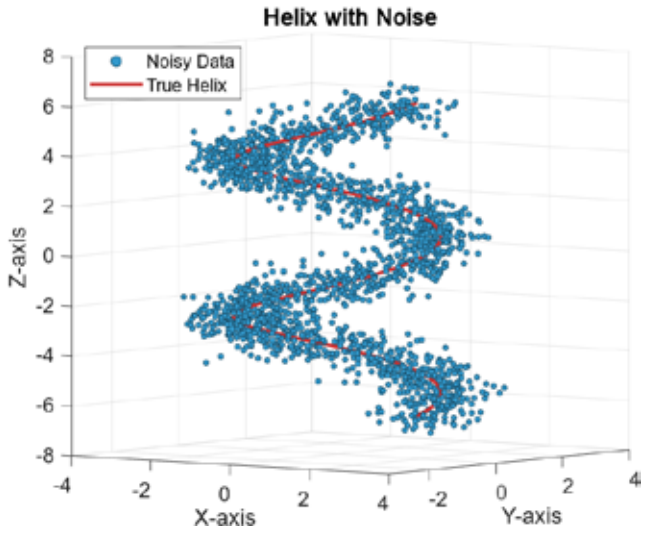


Figure 8: The datapoints of a dataset make up some helix-shape.

Now suppose that we have an extra feature for this dataset that is just a random variable. Think of this as a feature that has no further correlation with the other features.

However, this extra dimension will only contain noise and hence, the general structure of the dataset is preserved. But it might make it harder for t-SNE to make sense of the data. Note that it must be important in our case that t-SNE operates well under these circumstances since **there might be features in our client database that do not indicate if a client will leave**. Do you have an idea about how the 3D-representation of the changed dataset will be?

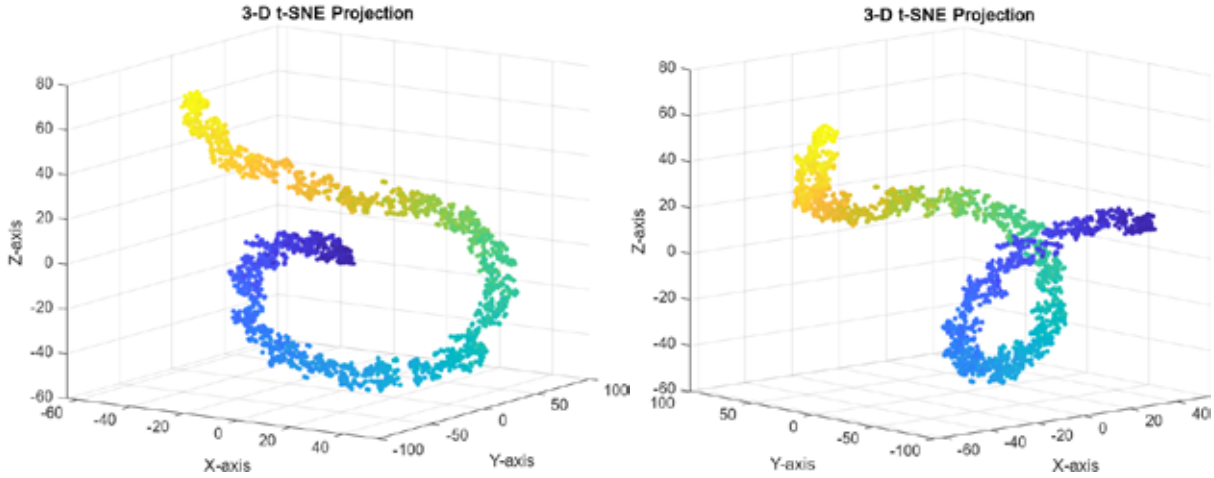


Figure 9: the 3D-result after applying t-SNE.

Note that the colour represents the original position of the datapoint in the curve. Hence, it is clear that t-SNE does a good job at recognizing that there is a line of datapoints for the helix.

Note that without background knowledge about the dataset, we have no intuition about the result we will get. We can only form conclusions based on the result t-SNE gives us.

Now suppose t-SNE gives the following 2D-result for our client dataset.

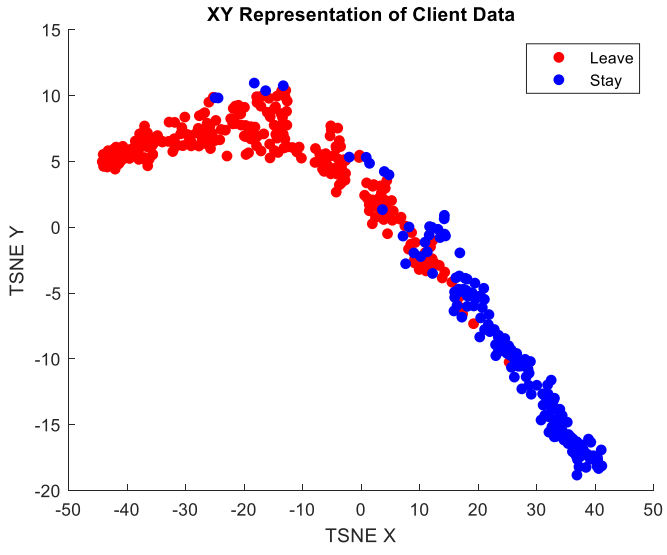


Figure 10: t-SNE output based on client dataset.

Based on this representation, it seems that a neural network or SVM might have an easy time in distinguishing leaving clients from staying clients. Hence, it might be easy to find how we can make clients stay. This is nice because now our boss can be assured that the problem can be solved given the client information present in the database.

Now assume that after applying t-SNE on the client dataset, the result would be the following:

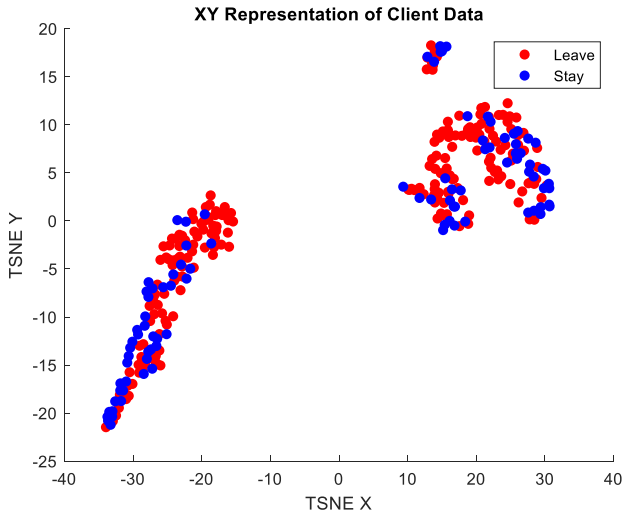


Figure 11: t-SNE output based on a different client dataset.

Since there is no great distinction between leaving or staying clients, it might be more difficult to find out how service can be improved. Hence, this might be an indication that the present client information is not sufficient to reliably estimate whether clients will leave or stay.

CONCLUDING t-SNE

In this article, it was observed that for **high-dimensional data it is difficult to get a feel for how it actually behaves**. This might make any objective in data science hard. A possible option is using algorithms that transform higher dimension data to a lower dimension. **The t-SNE algorithm is a good option for this, capturing the local and some global information by matching a lower-dimension distribution to a higher one**.

There are still some things that were left out. For example, how t-SNE exactly updates the position of the points in the lower-dimensional representation. Under the hood t-SNE will try to make sure that the distribution between points in the higher-dimension will align with the distribution of points in the lower-level embedding.

An alternative to t-SNE is UMAP. UMAP can be generally applied and deserves an article on its own. UMAP can be faster than t-SNE and often captures more of the global relationships, which can help in understanding the broader patterns in a complex dataset.

Sources:
https://www.youtube.com/watch?v=NEaUSP4YerM&ab_channel=StatQuestwithJoshStarmer
https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf
https://medium.com/@aeonaten/understanding-umap-uniform-manifold-approximation-and-projection-cede51c477d9
https://towardsdatascience.com/t-sne-clearly-explained-d84c537f53a
https://distill.pub/2016/misread-tsne/

*Dit is een gastbijdrage van **Jasper Visterin** voor SAI Update. Dit artikel schreef hij in het kader van zijn master-opleiding als data-ingenieur, en meer bepaald burgerlijk ingenieur wiskundige ingenieurstechnieken aan de KU Leuven.*

VOLGENDE EVENTS VOOR SAI.BE

2 OKT 2024
Partner event
CYBERSECURITY DAYS 2024
*Meer info vind je **hier***

8 OKT 2024
Partner event
BE CYBER
*Meer info vind je **hier***

15 OKT 2024
Webinar
13.00 – 14.00 uur
PRIVACY IN PRACTICE WITH SMART PSEUDONYMIZATION
*Meer info vind je **hier***

23 OKT 2024
Partner event
ANNACON 0X7E8
Provinciehuis Antwerpen
*Meer info vind je **hier***

25 OKT 2024
Online
13.00 – 14.00 uur
IT SUSTAINABILITY
*Meer info vind je **hier***

12 NOV 2024
Online
13.00 – 14.00 uur
PERSONEELSINTEGRITEIT VANUIT EEN SECURITY/ COMPLIANCY PERSPECTIEF
*Meer info vind je **hier***

18 NOV 2024
Online
13.00 – 14.00 uur
NIS2 ÉÉN MAAND LATER
*Meer info vind je **hier***

12 DEC 2024
19.30 – 21.30 uur
TECHNOVISION 2025: IT-TRENDS THAT WILL BECOME MAINSTREAM IN 2025
Avondconferentie
Van der Valk Hotel Brussels Airport (Brussel)
*Meer info vind je **hier***

Voor nieuwe events: neem zeker een kijkje op
WWW.SAI.BE

Beluister ***hier*** ook de extra edities van **SAI.NT, de podcast van SAI.BE.** De SAI.NT-podcast bespreekt de laatste thema's en actualiteit binnen het IT- en technologielandschap.

COLOFON

Werkten mee aan dit nummer: William Visterin (coördinatie), Robin Van den Bogaert, Stef Gyssels, Jasper Visterin en Marc Vael.

De missie van SAI.BE is om actuele en relevante IT kennis te delen op een objectieve en kwalitatieve manier met alle informatici in Vlaanderen en Brussel

Architecture

- Enterprise Architecture
- Data Architecture
- Business Capability Mapping
- BPMN / UML
- Decision Modelling
- IT Governance
- Distributed Systems
- Microservices Architecture

Cloud & Infrastructure

- Smart Cloud Deployment
- CASB
- Kubernetes
- Automation & Orchestration
- Open Service Mesh

Security

- Applied DevSecOps
- Zero Trust Implementation
- Cybersecurity Compliance Certifications
- Trusted Digital Identity Management
- Adaptive Security Awareness
- Intelligent Privacy Governance Automation
- SAST / SCA

Software Development

- Knowledge Graphs
- API management
- Low Code / No Code
- Composable Applications
- Customer Experience (CX)
- Code Quality

Data & AI

- Data Quality Management Tools
- Predictive Analytics Tools
- Data Scraping & Data Tracking
- Advanced Data Visualisation
- Regression-based Forecasting
- Machine Learning
- Data & AI Ethics By Design

Emerging Tech & Topics

- VR / AR / MR / XR
- Intelligent Process Automation
- Federated Learning
- The Autonomous Enterprise
- Sustainable IT
- Blockchain / NFT
- Quantum Computing
- The Future of Work

- ✓ SAI.BE begeleidt duizenden informatici **sinds 1967** doorheen een immer wijzigend IT landschap
- ✓ SAI.BE organiseert jaarlijks **gemiddeld 50 events**, waaronder avondconferenties, workshops, webinars, focus-meetings, speciale events en ook podcasts

SAI.BE publiceert elk kwartaal **het tijdschrift "SAI Update"** voor informatici, IT-experten, en IT beslissings-makers

MEER WETEN OF LID WORDEN?
Ga naar **www.sai.be/pagina/lidmaatschap/**

NEEM CONTACT OP
voorzitter@sai.be