Digital

VEJLEDNING

Dataklassifikation

DI Digital
1787 København V.
3377 3377
digital.di.dk
digital@di.dk

DI Digital – et branchefællesskab i Dansk Industri for virksomheder inden for it, tele, elektronik og kommunikation

Udgivet af: DI Digital

Redaktion: Henning Mortensen

ISBN: 978-87-7144-088-1

0.09.16

Digital

VEJLEDNING

Dataklassification

BAGGRUND

Trusselsbilledet for virksomheder er under konstant udvikling i disse år. Antallet af kendte sårbarheder er stigende år for år, og det samme er kompleksiteten og de indbydes afhængigheder mellem it-systemer. Antallet af kritiske og ukendte sårbarheder er også stigende. De it-kriminelle bliver dygtigere og dygtigere og finder til stadighed nye måder at angribe virksomhederne på. De optimerer og sætter ind, der hvor der er flest penge at hente i forhold til deres arbejdsindsats. Derfor er det ikke kun store virksomheder, der er deres mål. De mindre virksomheder, som typisk ikke har dedikerede ressourcer til at håndtere informationssikkerheden, er oplagte mål for simple men meget effektive angreb, der har små omkostninger for de it-kriminelle, men kan koste de mindre virksomheder dyrt. Derfor er der også stigende udfordringer med at vurdere, hvilken risiko man står overfor: Et år sker der måske stort set ingen vellykkede sikkerhedshændelser, og det næste år er virksomheden måske både udsat for tyveri via CEO-fraud og afpresning via ransomware - med betydelige tab til følge.

Virksomheden har selv kontrol over, hvilke digitale aktiver den er i besiddelse af. Derfor er det et godt – og afgørende – udgangspunkt for sikkerhedsarbejdet at identificere disse digitale aktiver og opdele dem i kategorier afhængigt af, hvilken værdi de har for virksomheden, hvor følsomme de er for virksomheden og efter lovmæssige krav. Denne øvelse kaldes dataklassifikation.

Helt overordnet har klassifikationen til formål at fastlægge, hvordan mennesker og systemer skal håndtere og beskytte data.

Hvis man ikke har klassificeret sine data, fremstår de alle som lige vigtige (eller ligegyldige) for virksomheden. Man beskytter alt lige godt (eller lige dårligt), og dermed kommer man typisk til at overinvestere i en form for sikkerhed, mens man underinvesterer i en anden form for sikkerhed. En anden fordel ved at klassificere sine data er, at man typisk opdager, at man er i besiddelse af data, som man ikke var klar over. Selv om dette ikke er genstanden for denne vejledning, skal det nævnes, at der ligger store forretningsmæssige potentialer i at kende sine data godt, således at man måske kan se nogle nye sammenhænge og tilbyde nye services eller finde et billede af, hvordan man fastholder kunder. Når man går i gang med arbejdet, bør målsætningen være, at man identificerer og klassificerer hver stump af data.

I denne vejledning vil vi give en kort og operationel tilgang til arbejdet med at klassificere sine data. Der vil blive præsenteret et forslag til, ud fra hvilke kriterier klassifikationen kan udformes I vejledningen vil det blive præsenteret, hvordan virksomheden bør have en proces for at identificere og klassificere sine data.

HVOR KOMMER DATA FRA?

Der sker en eksplosion af data i disse år. Faktisk viser estimater, at datamængden stiger med 40% hvert år. Mange af disse data har hele tiden været der, men i takt med at alting er blevet forbundet til internettet, som et internet af ting, IoT, stiger mængden. F.eks. har der hele tiden været data i vandmåleren. Men hvor man tidligere aflæste manuelt en gang hvert halve år, bliver der nu elektronisk aflæst løbende, i takt med at der forbruges. Tilsvarende kommer der data til og fra alle de andre digitale ting vi omgiver os med – fra målere, sensorer, skannere, osv.

På forbrugersiden er der indlejret kameraer som standard i pc'er, tablets, smartphones og smartTVs. De bruges til, at vi hele tiden deler billeder og optagelser med hinanden og dermed eksploderer datamængden fra den enkelte forbruger. Geolokationstjenester afgiver og modtager også data, så vi kan få de relevante tjenester, der hvor vi befinder os. De tjenester vi benytter os af, giver også data og skal fodres med data. Data kommer derfor f.eks. fra de terminaler der er nævnt ovenfor, wearables, digitale ure, sociale tjenester, osv.

Vi måler også på flere ting. Personbiler, lastbiler, containere og tankskibes position er ofte overvåget i vores data. Vores aktivitetsniveau og præsentationer, vores søvn vores adfærd, vores vægt og vores transaktioner måles. Det bidrager alt sammen med nye data.

Ud over disse spektakulære nye data, findes data naturligvis også der, hvor vi hele tiden har været vant til at adressere dem:

- i kontoradministrative systemer, som indeholder HR-data, løn, ferie og sygdomsdata
- i leverandørsystemer, med kommerciel information om bl.a. aftaler, fortrolighedserklæringer og priser
- i kundesystemer, som indeholder informationer om kunden, køb, tilbud, aftaler, licenser, konfiguration og login
- i drift og supportsystemer, som indeholder logdata, systembackup, supportsager m.v.
- i udviklingssystemer, som indeholder kommende generationer af det virksomheden skal tjene sine penge på.

Ovenpå alle disse kendte og nye data vil der også typisk være data om data, kaldet metadata. Systemer genererer disse data f.eks. for at det kan dokumenteres, hvornår en device har været i dialog med en anden device. Metadata bruges også for at kategorisere data af samme slags.

Det er ikke altid umiddelbart synligt, hvor data går hen. Mange data går direkte fra en elektronisk device til en anden uden menneskelig intervention. Devices reagerer intelligent på baggrund af de data de modtager fra andre devices.

Det er ikke alle data der er digitale og ligger fast i it-systemer. Man skal huske, at der stadig findes masser af data på papir, ligesom man skal huske at en del digitale data bliver lagt ud på medier som usb eller bliver lagt i cloud-tjenester.

Nogle data er også flygtige, og er kun at betragte som data i øjeblikket – f.eks. live videooptagelser. Men selv om data er flygtige, kan de godt påvirke sikkerheden.

Der er altså data overalt omkring os, og alt, hvad vi foretager os, skaber nye data. Der er altså langt flere data, end man lige umiddelbart går og tror, og dermed er vanskeligere end man lige umiddelbart tror at identificere, klassificere og beskytte disse data.

HVAD ER SIKKERHED EGENTLIG?

Når vi taler om at sikre data og systemer, handler det om at sikre tilgængelighed, fortrolighed og integritet samtidig.

Tilgængelighed betyder, at det skal være muligt for dem, der er autoriseret til det, at tilgå systemer og data, når de har brug for det.

Fortrolighed betyder, at kun autoriserede personer har ret til at tilgå systemer og data, og dermed at systemer og data ikke er tilgængelige for uvedkommende.

Integritet betyder, at data skal være pålidelige i betydningen komplette, korrekte og opdaterede, og der må altså ikke være nogen, som har haft mulighed for at manipulere med dem.

Når man skal klassificere data, bør det ske med udgangspunkt i disse tre begreber. Hvad betyder det for forretningen, at data via e-mailserveren ikke er tilgængelig i to timer eller 14 dage? Hvad betyder det for forretningen, at data fra udviklingsprojekter kopieres af uvedkommende – f.eks. en konkurrent? Hvilke konsekvenser har det for forretningen, hvis målerdata ændres som følge af en systemfejl?

Man kan ikke generelt sige noget om, hvorvidt tilgængelighed, fortrolighed eller integritet er vigtigst. I udgangspunktet er de tre begrebet sidestillet. Men når man konkret vurderer det enkelte datasæt ud fra de tre begreber, vil et af dem ofte være af større betydning end de andre. For forskningsdata vil fortrolighed måske være vigtigere end tilgængelighed, fordi det er afgørende, at konkurrenter ikke får adgang til den udvikling, som virksomheden skal bygge sin fremtid på, hvorimod det betyder mindre, at projektet ligger stille en dag. Andre data giver måske ikke mening isoleret set, hvis de slipper ud, hvorfor fortroligheden betyder mindre, men til gengæld kan de være afgørende for driften af virksomhedens produktion og derfor er disse datas tilgængelighed afgørende. Betydningen af de tre begreber kan i øvrigt ændres over tid, således at data, som der er behov for høj fortrolighed om i dag ikke behøver at være omgærdet af fortrolighed i morgen.

EN MINDRE VIRKSOMHEDS TILGANG

Den mindre virksomhed har typisk ikke ressourcer og kompetencer til at iværksætte en fuld dataklassifikation ud fra de tre begreber, fortrolighed, tilgængelighed og integritet. I dette tilfælde må man gå intuitivt til værks og "forfatteren" af et dokument eller et sæt data må fra gang til gang vurdere konsekvenserne ved at data kommer til uvedkommendes kendskab, uautoriseret ændres eller ikke er tilgængeligt. På den baggrund kan man så opdele data i offentlige, interne eller fortrolige.

For hver kategori kan man så opstille nogle leveregler for data, f.eks.:

Offentlige data
Data må sendes til kunder, samarbejdspartnere og hvem der i øvrigt kunne være

interesseret. De må sendes som almindelig brev eller e-mail. Data af denne type kan opbevare, hvor man i afdelingen har aftalt det er mest praktisk.

Interne data
Data må kun deles med relevante samarbejdspartnere og med ansatte i virksomheden. De må sendes som almindelig brev, men skal de sendes med e-mail skal det ske krypteret. Data af denne type kan opbevares, hvor man i afdelingen har aftalt det er mest praktisk. Den enkelte medarbejder skal udvise fortrolighed omkring data/informationen.

Fortrolige data
Data må ikke sendes til kunder og samarbejdspartnere, medmindre det er godkendt af en direktør. Internt skal de kun være tilgængelige for dem, det er relevant for. Skal data sendes med brev skal det være anbefalet, og hvis de sendes med e-mail skal det være krypteret. Data må kun opbevares på på forhånd definerede steder. Den enkelte medarbejder skal udvise fortrolighed omkring data/informationen.

SKABELON FOR KLASSIFIKATION

De større virksomheder bør gennemføre en fuld dataklassifikation med udgangspunkt i tilgængelighed, fortrolighed on integritet. Der findes ganske mange forskellige bud på, hvordan data kan klassificeres. Udgangspunktet for klassifikationen er som regel, hvilken skade det vil forvolde virksomheden, hvis data afsløres for uvedkommende, hvorfor klassifikationen ensidigt fokuserer på fortrolighed og ikke på tilgængelighed og integritet.

Som det gerne skulle være tydeligt af ovenstående, er det imidlertid ikke tilstrækkeligt at fokusere på fortrolighed. Derfor skal man for at få en præcis klassifikation fokusere på alle tre sikkerhedsparametre og lave tre skabeloner.

Skabelon for fortrolighed

Der findes to gode eksempler på klassifikation af data, hvor der alene er fokus på fortrolighed, og ikke på tilgængelighed og integritet. Det ene stammer fra kontrol 8.2.1 i sikkerhedsstandarden iso27002[1]. Den anden stammer fra statsministeriets sikkerhedscirkulære[2] og anvendes især til at beskytte data relateret til EU og NATO. Vi har gengivet disse klassifikationer i bilag A. Disse to terminologier kan sammenfattes til en klassifikation af fortrolige oplysninger, som er operationel:

- **Almindelige oplysninger**
  Informationer, hvor offentliggørelse er naturlig eller ikke forvolder nogen skade for virksomheden eller virksomhedens relationer.

  (I denne klasse findes f.eks. informationer om produkter og services, salgsmateriale, nyhedsbreve, tekniske opdateringer, manualer og installationsvejledninger)

[1] http://www.iso.org/iso/catalogue_detail?csnumber=54533.
[2] https://www.retsinformation.dk/forms/R0710.aspx?id=166206.

- **Fortrolige oplysninger**
  Informationer, som er forbeholdt virksomheden og evt. dens partnere, og hvor offentliggørelsen kan forårsage mindre væsentlige problemer eller gener for virksomheden eller virksomhedens relationer.

  (I denne klasse findes f.eks. beskrivelse af visse interne forretningsgange, salgspriser, lister over kunder, lister over licenser, kundespecifikke mails, de fleste IoT-data, kontrakter og aftaler)

- **Hemmelige oplysninger**
  Informationer, som er forbeholdt en afgrænset del af virksomhedens personale og eventuelt en afgrænset del af en partners personale, og hvor offentliggørelse kan forårsage væsentlige problemer med konsekvenser for driften eller taktiske målsætninger.

  (I denne klasse findes f.eks. oplysninger om kundernes infrastruktur, logningsdata, sikkerhedsanalyser, HR-data og andre personoplysninger, eller andre data omfattet af lovgivning, ledelsesnotater, data tilknyttet udviklingsprojekter)

- **Tophemmelige oplysninger**
  Informationer, som er forbeholdt virksomhedens topledelse, samt få udvalgte specialister i virksomheden og eventuelt undtagelsesvist en afgrænset del af partneres personale, som præciseret i kontrakt eller som følge af lov, og hvor offentliggørelse kan forårsage alvorlige problemer for driften eller strategiske målsætninger eller decideret sætter virksomhedens overlevelse på spil.

  (I denne klasse findes f.eks. passwords til kundernes systemer, ledelsesstrategiske beslutninger om forretningens udvikling, data tilknyttet udviklingsprojekter af væsentlig betydning for virksomhedens fremtid (f.eks. med henblik på at udtage patent)).

Skabelon for tilgængelighed

Udgangspunktet for at klassificere datas tilgængelighed er, hvor lang tid virksomheden kan leve med ikke at have adgang til et givent sæt data. Dette afhænger i høj grad af, om man kan udføre sit arbejde på en anden måde end ved at have adgang til data – f.eks. via manuelle processer.

- **Høj tilgængelighed**
  Det er afgørende for forretningens virke, at der er adgang til data og arbejdet kan ikke erstattes af manuelle processer.

  (I denne klasse findes f.eks. data som er afgørende for at holde produktionen kørende, eller for at online salg i e-handelsvirksomheder kan gennemføres)

- **Medium tilgængelighed**
  Tilgængelighed har betydning for forretningen, men visse funktioner kan udføres manuelt eller udskydes i en vis periode, indtil der igen skabes tilgængelighed.

  (I denne klasse findes f.eks. faktureringssystemer, som i en kortere periode

kan gennemføres via manuelle processer eller besvarelse af spørgsmål, som kan foretages telefonisk i stedet for via e-mail)

- **Lav tilgængelighed**
  Tilgængeligheden har mindre betydning for forretningen, og de vigtigste funktioner kan gennemføres i en længere periode uden adgang til data.

  (I denne klasse findes f.eks. adgang til fysiske arkiver)

Skabelon for integritet

Udgangspunktet for at klassificere datas integritet er, hvor præcise data skal være, for at virksomheden ikke laver væsentlige fejl. Dette afhænger f.eks. af, om data inden for et givent interval er tilstrækkeligt til, at virksomheden kan træffe de rette beslutninger, eller om data til enhver tid skal være meget præcise, for at beslutninger bliver korrekte.

- **Høj integritet**
  Forretningskritiske beslutninger bliver taget på grundlag af data.

  (I denne klasse findes f.eks. data der bliver brugt til ledelsesbeslutninger, data som grundlag for fakturaer eller data som anvendes til at bestemme kvaliteten af et produkt)

- **Medium integritet**
  Data anvendes til beslutninger, der ikke er kritiske for forretningen.

  (I denne klasse findes f.eks. data, som skaber økonomisk overblik, data til auditering eller efterfølgende kvalitetskontrol)

- **Lav integritet**
  Data anvendes aldrig til forretningskritiske beslutninger.

  (I denne klasse findes f.eks. data om foreninger tilknyttet virksomheden, generel rengøringsservice, frokostordninger og lignende)

PROCES I VIRKSOMHEDEN

At klassificere data er ikke noget, man kan gøre en gang for alle. Virksomheden bør fastlægge en proces for vedvarende klassifikation af data.

Den sikkerhedsansvarlige har en væsentlig rolle at spille i arbejdet med klassifikationen:

- Den sikkerhedsansvarlige bør først få godkendt klassifikationen af ledelsen.
- Derefter bør dataklassifikation skrives ind i sikkerhedshåndbogen, som en retningslinje.
- Yderligere bør den sikkerhedsansvarlige kommunikere klassifikationen ud i organisationen.
- Videre bør den sikkerhedsansvarlige få fastlagt ejerskab til data ude i organisationen og sikre, at ejerne foretager en initial klassifikation. Vurderingen skal tage udgangspunkt i tilgængelighed, fortrolighed og integritet samt risikoen

ved, at disse ikke er på plads. Det er vigtigt, at man kommer hele vejen rundt om de mange kilder til data, som virksomheden måtte have.
- Dernæst godkendes klassifikationen af den dataansvarlige.
- Endelig bør der fastlægges en politik for, i hvilket omfang eksisterende data skal klassificeres.

Typisk har den sikkerhedsansvarlige ikke overblik over, hvilke data der findes i organisationen. For at identificere og klassificere data må der derfor søges input fra de forretningsansvarlige rundt omkring i organisationen. Den sikkerhedsansvarlige kan typisk bede virksomhedens forskellige chefer redegøre for:

- Hvilke kategorier af data behandles i din funktion (f.eks. kundedata, HR-data, data i et udviklingsprojekt)?
- Hvor stammer de pågældende kategorier af data fra (f.eks. fra IoT-udstyr, indkøbt fra tredjemand, indsamlet som et led af forretningen)?
- Hvor går de pågældende kategorier af data hen (f.eks. overførsel til eksterne)?
- Med udgangspunkt i tilgængelighed, fortrolighed og integritet hvor afgørende er hver kategori af data så for forretningen (f.eks. værdi og følsomhed)?
- Er kategorierne af data omfattet af lovgivning (f.eks. personoplysninger)?

På baggrund af denne proces skabes et centralt overblik over data, som kan bruges til både dataklassifikationen og til det videre sikkerhedsarbejde. Den sikkerhedsansvarlige skal efterfølgende vurdere, om han er enig i den foreslåede klassifikation.

Efterfølgende skal der tages initiativ til at markere data med klassifikation fremadrettet. På en række områder kan der installeres software, som sikrer, at den, som behandler data, tager stilling til klassifikationen, når data skabes. Der kan også på enkelte områder installeres software, som identificerer forskellige syntakser i data og klassificerer dem på den baggrund – f.eks. syntaksen for et CPR-nummer.

Klassifikation af de eksisterende data kan være en uoverkommelig byrde. Det foreslås derfor, at disse data alene klassificeres i takt med, at de tilgås.

VIDERE SKRIDT

På baggrund af klassifikationen af data kan virksomheden gå i gang med at iværksætte beskyttelse af de enkelte data. Dette er ikke genstanden for denne vejledning. Men i overordnede termer skal virksomheden:

- Lave en risikovurdering af, hvilke risici de enkelte data eller klasser af data står overfor.
- Identificere hvilke sikkerhedstiltag, der er i forvejen er iværksat. Sikkerhedstiltag skal forstås bredt som menneskelig awareness om sikkerhed, processer for håndtering af aktiverne og som tekniske tiltag. Man kan tage udgangspunkt i ledelsesværktøjet for styring af sikkerheden i ISO27001 og sikkerhedskontrollerne i ISO27002
- Fastlægge hvilke supplerende sikkerhedstiltag, der eventuelt skal suppleres med
- Overveje, hvordan virksomheden vil håndtere restrisikoen – f.eks. gennem forsikring.

TJEKLISTE

Klassifikationsterminologi
Mindre virksomheder, som ikke kan gennemføre en egentlig gennemgang af deres data og en efterfølgende dataklassifikation, bør intuitivt klassificere deres data i offentlige, interne og fortrolige.

Større virksomheder bør klassificere sine data på baggrund af en prædefineret dataklassifikation. Dataklassifikationen kan ud fra en vurdering af fortrolighed, tilgængelighed og integritet antage formen:

Fortrolighed	Tilgængelighed	Integritet
• Almindelige	• Høj	• Høj
• Fortrolige	• Medium	• Medium
• Hemmelige	• Lav	• Lav
• Tophemmelige		

Klassifikationsproces
Virksomheden bør iværksætte en proces således, at data som minimum fremadrettet klassificeres:

- Ledelsesgodkendt klassifikationen.
- Klassifikationen gøres tilgængelig for medarbejderne i sikkerhedshåndbogen.
- Klassifikationen kommunikeres ud i organisationen.
- Der fastlægges ejerskab til data i organisationen og ejerne forestår initial klassifikation baseret på tilgængelighed, fortrolighed og integritet.
- Den sikkerhedsansvarlige godkender den konkrete klassifikation af data.
- Eksisterende data klassificeres i takt med at de anvendes.

Input til videre sikkerhedsarbejde
Klassifikationen kan anvendes i det videre sikkerhedsarbejde til:

- Risikovurdering
- Vurdering af eksisterende sikkerhedsinitiativer og evt. supplerende sikkerhedstiltag
- Håndtering af restrisiko

BILAG A

Der findes to gode klassiske eksempler på klassifikation af data, hvor der alene er fokus på fortrolighed, og ikke på tilgængelighed og integritet. Disse er gengivet nedenfor.

For det første er dataklassifikation med udgangspunkt i fortrolighed nævnt som kontrol 8.2.1 i sikkerhedsstandarden iso27002[3]. Her er klassifikationssystemet baseret på fire niveauer, som er unavngivne:

- Offentliggørelse forårsager ingen skade
- Offentliggørelse forårsager mindre væsentlige problemer for driften
- Offentliggørelsen har en væsentlig kortsigtet indvirkning på driften eller taktiske målsætninger
- Offentliggørelse har en alvorlig indvirkning på langsigtede strategiske målsætninger eller sætter organisationens overlevelse på spil

For det andet er dataklassifikationen med udgangspunkt i fortrolighed nævnt i sikkerhedscirkulæret[4] og anvendes især til at beskytte data relateret til EU og NATO. Også her er der fire klassifikationsniveauer, som dog er navngivne:

- TIL TJENESTEBRUG (»NATO RESTRICTED«, »RESTREINT UE« / »EU RESTRICTED«). Denne klassifikationsgrad anvendes om informationer, der ikke må offentliggøres eller komme til uvedkommendes kendskab.
- FORTROLIGT (»NATO CONFIDENTIAL«, »CONFIDENTIEL UE« / »EU CONFIDENTIAL«). Denne klassifikationsgrad skal anvendes om informationer, hvis videregivelse uden dertil indhentet bemyndigelse ville kunne forvolde Danmark eller landene i NATO eller EU skade.
- HEMMELIGT (»NATO SECRET«, »SECRET UE« / »EU SECRET«). Denne klassifikationsgrad skal anvendes om informationer, hvis videregivelse uden dertil indhentet bemyndigelse ville kunne forvolde Danmark eller landene i NATO eller EU alvorlig skade.
- YDERST HEMMELIGT (»COSMIC TOP SECRET«, »TRÈS SECRET UE« / »EU TOP SECRET«). Denne klassifikationsgrad skal anvendes om informationer, hvis videregivelse uden dertil indhentet bemyndigelse ville kunne forvolde Danmark eller landene i NATO eller EU overordentlig alvorlig skade.

[3] http://www.iso.org/iso/catalogue_detail?csnumber=54533.
[4] https://www.retsinformation.dk/forms/R0710.aspx?id=166206.