Autonomi i cyberoperasjonar

Mass Soldal Lund

Høgskolen i Innlandet, Institutt for organisasjon, styring, leiing, Rena
Forsvarets høgskole/Cyberingeniørskolen, Lillehammer

Innleiing

Fyrste setning i forordet til antologien *Adaptive Autonomous Secure Cyber Systems* lyder: «Autonomy in physical and cyber systems promises to revolutionize cyber operations.»[1] Dette er symptomatisk for spørsmålet om autonomi i cyberoperasjonar: Det finst ei stor forventing om at autonomi, og ikkje minst intelligente system, vil ha mykje å seie for framtidas cyberoperasjonar. Men sjølv om det er mogleg å snakke om ein viss grad av autonomi i cyberoperasjonar, er dette løftet førebels ikkje oppfylt, og det er eit ope spørsmål i kva grad det vil bli det.

Dette kapittelet vil utforske autonomi i cyberoperasjonar. Målet er ikkje å skrive science fiction, men å freiste å gje eit bilete av stoda for autonome cyberoperasjonar i dag, og av kva ein kan forvente i nær framtid.[2] Kapittelet vil ta føre seg korleis ein kan forstå autonomi i cyberoperasjonar, med andre ord: *Kva er ein autonom cyberoperasjon?* Det vil vere konsentrert om dei fordelane autonomi i slike operasjonar vil kunne gje, og om kva som eventuelt vil vere baksider og «showstopparar», men det vil avgrense seg frå juridiske, personvernmessige og etiske sider ved cyberoperasjonar.

Det vil vere meiningslaust å snakke om autonomi i cyberoperasjonar utan å ta føre seg programvare. Kapittelet vil difor starte med generelle tankar om autonomi – og dei relaterte eigenskapane automatikk og intelligens – i programvare, for deretter å sjå på kva dette, på generelt nivå, har å seie for cyberoperasjonar. Til slutt vil kapittelet gå konkret inn på offensive og defensive cyberoperasjonar, og sjå på kva rolle autonomi har å seie for slike operasjonar i dag.

[1] Sushil Jajodia mfl. (red.), *Adaptive Autonomous Secure Cyber Systems* (Cham: Springer, 2020), v.
[2] Fordi cyberoperasjonar ofte er gjennomført i det skjulte, kan det òg eksistere cyberoperasjonar i dag som fyrst vil verte kjente i framtida.

Referanse: Mass Soldal Lund, «Autonomi i cyberoperasjonar» i *Autonomi i militære operasjoner*, red. Camilla Guldahl Cooper og Dag Henriksen (Oslo: Cappelen Damm Akademisk, 2023), 139-160.

Programvare, automatikk, autonomi og intelligens

Framveksten av datateknologien som cyberdomenet og cyberoperasjonar kvilar på, er ei historie om automatisering av prosessar som tidlegare vart utført av menneske, anten det er snakk om talknusing, administrativ datahandsaming eller prosesstyring. Dette gjer at alle arbeidsprosessar som involverer datateknologi, i større eller mindre grad vil vere automatisert i den forstand at dei omfattar operasjonar som blir utført automatisk (utan menneskeleg inngripen) av programvare. Men når kan ein seie at ei programvare opererer autonomt? I introduksjonen til denne boka er autonomi skildra som ein funksjon ved eit system som gjer det i stand til å agere eller handle for å oppnå førehandsdefinerte mål og tilpasse utførsel av oppgåver basert på sensorisk input. Dette er ein måte å seie at autonomi er evna til å ta sjølvstendige avgjerder eller val (basert på input). Utfordringa er at ein slik definisjon i seg sjølv ikkje er nok til å skilje mellom automatikk og autonomi i programvare. Evna til å ta val basert på data er ein heilt grunnleggjande funksjon i koding eller programmering av datamaskiner og figurerer sentralt både i den historiske utviklinga og i definisjonar av moderne datateknologi.[3] Intuitivt vil ein difor ikkje forstå kvart av dei tusenvis av små vala koda inn i programvare som teikn på autonomi; autonomi er gjerne forstått som noko kvalitativt annleis eller «meir» enn (høg grad av) automatikk. Ofte vil ein assosiere autonomi med «store» eller «viktige» avgjerder (i motsetnad til dei små og trivielle som programvare gjer heile tida) eller med «smarte» eller «intelligente» system. Men kva som blir oppfatta som ikkje-trivielle avgjerder, og kva som blir oppfatta som intelligens i datasystem, er ikkje konstant; det endrar seg i tråd med den (data)teknologiske utviklinga. Eigenskapar ved datasystem som i dag blir sett som ordinære og trivielle, har ein gong i tida vore oppfatta som autonomi og intelligens.[4] Tilsvarande må ein ta høgde for at det ein i dag oppfattar som autonomi og intelligens, i framtida kan bli sett som ordinære og trivielle eigenskapar ved datasystem.

Konklusjonen må difor bli at autonomi er ein kvalitativ og kontekstavhengig eigenskap som ikkje let seg definere utan å ta omsyn til domene, oppgåver og teknologisk

[3] Sjå t.d. Thomas Haigh, Mark Priestley og Crispin Rope, *ENIAC in Action: Making and Remaking the Modern Computer* (Cambridge, MA: MIT Press, 2016), 231–257.

[4] Eit illustrerande eksempel er følgjande sitat av Edmund C. Berkeley frå boka *Giant brains or machines that think* frå 1949, der han skriv om dei fyrste programmerbare reknemaskinene, forlauparar til moderne datamaskiner: «A machine can handle information; it can calculate, conclude, and choose; it can perform reasonable operations with information. A machine, therefore, can think.»; Edmund C. Berkeley, *Giant Brains or Machines that Think* (New York: John Wiley & Sons, 1949), 5; sjå også Mass Soldal Lund, «Menneskemaskina», *Syn og segn* 126, nr. 1 (2020).

state-of-the-art. Det følgjande vil difor ikkje freiste å gje generelle definisjonar av autonomi, men heller forsøkje å fange det ein kan forstå som *autonomi i cyberoperasjonar i dag*.

Autonomi i cyberoperasjonar

Cyberoperasjonar er utnytting av datasystem for å få dei til å utføre oppgåver som ligg utanfor intensjonen eller hensikta med systema, og eit strev for å oppretthalde intensjonen (normal funksjon) i systema i møte med slik utnytting. I det fyrste tilfellet snakkar vi om offensive cyberoperasjonar, medan det andre tilfellet er det vi kallar defensive cyberoperasjonar.[5] Trass likskapen i namn er offensive og defensive cyberoperasjonar relativt ulike aktivitetar og vil difor bli handsama separat.[6] Men fyrst kan det vere greitt å sjå på nokre sider ved cyberoperasjonar som er felles for dei offensive og dei defensive operasjonane.

For det fyrste er cyberoperasjonar (i all hovudsak) utførte av menneske – som ein kan omtale som operatørar – med hjelp av programvare. Sjølv om det finst unntak[7], vil dei vere mindre interessante, sidan spørsmålet om autonomi vil handle om relasjon mellom programvara og den menneskelege operatøren.

For det andre er cyberoperasjonar styrte av prosedyrar. Slike prosedyrar kan vere formelle prosessar, beskriven beste praksis eller etablerte arbeidsmetodikkar og kan såleis vere meir eller mindre formaliserte. Det er openbert at dei vil variere mellom organisasjonar, men det viktige her er at dei kan generaliserast til idealiserte sekvensar av steg som skildrar typiske offensive og defensive cyberoperasjonar frå ståstaden til operatørane. Detaljane i slike idealiserte prosedyrar vil nedanfor bli gjeve for offensive og defensive operasjonar kvar for seg, men eksistensen av dei gjev høve til å skildre autonomi i cyberoperasjonar meir presist: *programvaras evne i ein cyberoperasjon til å ta avgjerder og utføre neste steg i prosedyren utan inngripen frå den menneskelege operatøren*. Ein slik definisjon gjev òg ei moglegheit til å operere med gradar av autonomi. For kvart steg i prosedyren vil graden av autonomi vere

[5] Mass Soldal Lund, «Cyber som operasjonsdomene», *Norsk militært tidsskrift* 186, nr. 1 (2017): 30–31.
[6] Dette er i tråd med det relativt strenge skiljet mellom offensive og defensive cyberoperasjonar ein finn i norsk og Nato-doktrine, der det er typen verkemiddel som avgjer om ein cyberoperasjon er offensiv eller defensiv, sjå Forsvaret, *Forsvarets fellesoperative doktrine* (Oslo: Forsvarsstaben, 2019), 125–126; Nato, *Allied Joint Doctrine for Cyberspace Operations*, Allied Joint Publication 3.20, Utg. A, Versjon 1 (Genève: Nato Standardization Office, 2020), 4, 16–17; dette skil seg frå til dømes amerikansk doktrine, som tillèt offensive verkemiddel i operasjonar med defensive målsettingar, sjå Joint Chiefs of Staff, *Cyberspace operations*, Joint Publications 3-12 (Washington, D.C.: Joint Chiefs of Staff, 2018), II-2–II-9, https://irp.fas.org/doddir/dod/jp3_12.pdf
[7] Eksempel kan vere ein offensiv operasjon der operatøren gjettar innloggingsinformasjon og utelukkande utnyttar eit datasystem gjennom tilgangen det gjev, eller ein defensiv operasjon der operatøren utelukkande gjer manuell analyse av systemloggar og manuelle systemendringar.

gjeve av i kor stor grad den menneskelege operatøren kan monitorere og moderere programvaras handlingar, medan for operasjonen som heilskap vil graden av autonomi vere summen av autonomi i einskildstega.[8]

Med ei slik forståing blir «intelligens» og autonomi uavhengige storleikar; ein kan ha både «dum» programvare som tek avgjerder basert på enkle kriterium, men like fullt utan menneskeleg involvering, og «intelligent» programvare som gjer avanserte analysar (basert på til dømes maskinlæringsmetodar), men som likevel presenterer dei til ein menneskeleg operatør for avgjerder. Potensialet for bruk av kunstig intelligens og maskinlæring blir drøfta spesifikt for offensive og defensive cyberoperasjonar nedanfor, men ein generell observasjon er at i andre domene har slike metodar vist seg nyttige og effektive for lukka problem (spesifikke oppgåver i ein avgrensa og veldefinert omgjevnad), men mindre vellykka for opne problem.[9] Eit sentralt spørsmål blir difor i kva grad cyberoperasjonar høyrer heime i ei lukka eller ei open verd.

Denne boka handlar om autonomi i militære operasjonar, og dette kapittelet vil såleis vinkle seg inn mot militære cyberoperasjonar. Når det er sagt, er det ikkje naudsynleg så stor skilnad på militære cyberoperasjonar og cyberoperasjonar gjennomførte av ikkje-militære organisasjonar; verktøy, teknikkar og taktikkar vil vere samanliknbare, og det er heller ikkje alltid eit skarpt skilje mellom militære og sivile aktørar i cyberdomenet.[10] Det avgjerande skiljet mellom militære og ikkje-militære cyberoperasjonar vil difor ikkje vere metodane, men heller dei organisatoriske og formelle rammene og kva slags system som er mål for operasjonane. Som eit eksempel vil offensive cyberoperasjonar i vestleg doktrine (til dømes Noreg, USA, Nato) vere underlagt formelle prosessar for fellesoperativ målutveljing på same måte som andre militære verkemiddel.[11] Dette legg (til dømes folkerettslege) avgrensingar på bruken av verkemidla som ikkje-statlege eller ikkje-vestlege aktørar ikkje naudsynleg vil halde seg med.[12] Den militære konteksten vil difor ikkje påverke dei praktiske og teoretiske

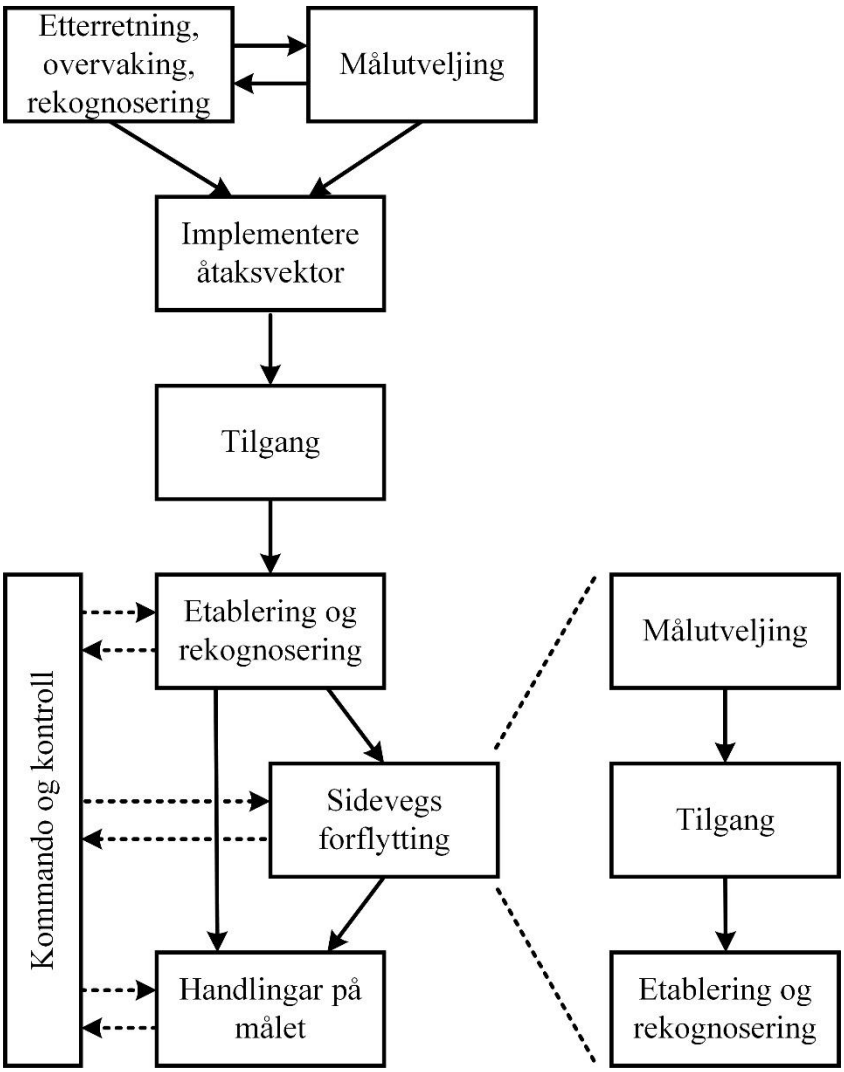
[8] Gregory Conti og David Raymond, *On Cyber. Towards an Operational Art for Cyber Conflict* (Kopidion Press, 2017), 220–224.
[9] Gary Marcus og Ernest Davis, *Rebooting AI. Building Artificial Intelligence We Can Trust* (New York: Pantheon Books, 2019), 113–114.
[10] Conti og Raymond, *On Cyber*, 12–15; Mass Soldal Lund, «CND-konsept for taktiske nettverk», i *Ledelse i Cyberdomenet*, red. Øyvind Jøsok og Benjamin J. Knox (Lillehammer: Forsvarets ingeniørhøgskole, 2018), 38–42; Jan Terje Ringstad, *Mind the Gap – An Exploratory Study of Commercial and Military Computer Security Incident Response Teams (CSIRTs) – Are Incident Response (IR) and Computer Security Incident Response Teams (CSIRTs) Forensic Ready in the Information Domain?*, Masteroppgåve (NTNU, 2017).
[11] Conti og Raymond, *On Cyber*, 179–212; Joint Chiefs of Staff, *Cyberspace operations*, I-8; Forsvaret, *Forsvarets fellesoperative doktrine*, 130–131; Nato, *Allied joint doctrine for cyberspace operations*, 27.
[12] Mass Soldal Lund, «Cybertrusselen og elektronisk krigføring – kva kan vi sjå føre oss i gråsona?», *Norsk militært tidsskrift* 191, nr. 3 (2021).

sidene ved autonomi i cyberoperasjonar i særleg grad, men vil kunne ha ein del å seie for kva ein reknar som fordelar og baksider ved autonomi.

Offensive cyberoperasjonar

Ein offensiv cyberoperasjon vil ofte vere avhengig av tilgang til kunnskap (til dømes om sikkerheitshol) og teknikkar som forsvararane av målsystemet ikkje kjenner til. Difor vil detaljane om offensive cyberoperasjonar som regel vere godt bevarte løyndomar. Slik hemmeleghald vert forsterka av at offensive cyberoperasjonar ofte høyrer etterretningsverda til, og av etterretningsorganisasjonars hang til å halde arbeidsmetodane sine løynde. Denne mangelen på openheit gjer at modellar for offensive cyberoperasjonar gjerne er skildra med utgangspunkt i forsvararens observasjonar av offensive operasjonar. Døme på dette er den mykje siterte modellen *Intrusion Kill Chain* eller *Cyber Kill Chain*, utvikla av Lockheed Martin,[13] og *ATT&CK*, som er eit rammeverk og ein kunnskapsbase over kjente taktikkar og teknikkar for cyberåtak utvikla av tenketanken MITRE.[14] Grant mfl. presenterer ein idealisert modell som eksplisitt er tenkt for statlege organisasjonar. Modellen er ein syntese av modellar identifisert gjennom ein litteraturstudie; fleire av dei er skildra frå åtakarens synsvinkel, men langt ifrå alle frå ein statleg ståstad.[15] Eit føredrag frå NSA Tailored Access Operations gjev også eit innblikk i offensive cyberoperasjonar.[16] Ein modell for offensive cyberoperasjonar basert på desse kjeldene er vist i Figur 1.

[13] Eric M. Hutchins, Michael J. Cloppert og Rohan M. Amin, «Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains», i *Proceedings of the 6th International Conference on Information Warfare and Security*, red. Leigh Armistead (Reading, Storbritannia: Academic Publishing International, 2011); Lockheed Martin, «The Cyber Kill Chain», lese 05.09.2022, https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html.
[14] Mitre, «Att&ck», lese 03.08.2022, https://attack.mitre.org/
[15] Tim Grant, Ivan Bruke og Renier van Heerden, «Comparing Models of Offensive Cyber Operations», i *Leading Issues in Cyber Warfare & Security Research*, Bd.2, red. Julie Ryan (Reading, Storbritannia: Academic Conferences and Publishing International, 2015).
[16] Rob Joyce, «Disrupting nation state hackers», Video, 34:55, *USENIX Enigma*, 27.01.2016, https://www.usenix.org/conference/enigma2016/conference-program/presentation/joyce



Figur 1 Offensive cyberoperasjonar

Operasjonen russiske hemmelege tenester utførte mot det demokratiske partiet i USA i 2016, kan illustrere prosessen. *Etterretning* gjennom opne kjelder gav åtakaren kunnskap om tilsette i Democratic Congressional Campaign Committee (DCCC). Som *åtaksvektor* nytta åtakaren phishing-e-post som narra ein tilsett i DCCC til å avsløre passordet sitt. Dette passordet gav åtakaren *tilgang* til datanettverket til DCCC. Med denne tilgangen *etablerte* han seg i nettverket ved å installere ei skadevare, som han så kunne styre gjennom ein *kommando- og kontrollkanal*. Ved å nytte skadevara til å overvake brukarar, identifiserte åtakaren ei VPN-kopling (ein kryptert nettverkstunnel), med tilhøyrande passord, frå DCCC til Democratic National Committee (DNC). Dette utnytta åtakaren til å *forflytte seg sidevegs* til datanettverket til DNC og til å etablere seg også der. Både i DCCC og i DNC nytta åtakaren skadevara til å søke etter interessante filer (med søkjetermar som «Hillary», «DNC», «Cruz»

og «Trump») som han samla og komprimerte før filene vart lasta ned til ein server som var leigd for føremålet.[17]

I eksempelet er kvart steg i prosessen handlingar utført av ein menneskeleg aktør, i mange av tilfella med støtte av programvare eller skadevare. Vidare involverer kvart steg avgjerdslar om vidare handlingar basert på innhenta data: phishing basert på kjennskap til brukarar, tilgang basert på kjennskap til passord, sidevegs forflytting basert på kjennskap til VPN-kopling mot DNC, uthenting av informasjon basert på søk. Fleire av desse handlingane føreset at åtakaren har tilgang gjennom internett til systemet og skadevara han planta.

Fordi stega i prosessen (som regel) vil vere støtta av programvare (inkludert skadevare), vil prosessen (som regel) ha ein viss grad av automatikk. Til dømes vil rekognosering av eit datanettverk ofte vere støtta av programvare som automatiserer søk etter maskiner og sårbarheiter i nettverket. Autonomi vil derimot seie at det er programvare som (basert på slike søk) tek avgjerder om neste steg i operasjonen, slik som avgjerder som kva maskiner som skal vere mål eller kva for åtaksvektor (sårbarheit) som skal utnyttast.

Eksempel på autonome offensive cyberoperasjonar

Dataormar er skadevare som spreiar seg av seg sjølv, og som dimed, per definisjon, har ein grad av autonomi. Dataormar kan difor sjåast som det fremste eksempelet på autonom skadevare, og ein autonom offensiv cyberoperasjon vil som regel nytte ein form for dataorm eller skadevare med «ormeliknande» eigenskapar. To illustrerande døme på offensive cyberoperasjonar som nytta dataormar, er det russiske åtaket på ukrainsk næringsliv med ormen NotPetya i 2017, og Operation Olympic Games, eit amerikansk-israelsk åtak mot eit uranopprikingsanlegg i Iran.

NotPetya-operasjonen starta med at åtakarane etablerte tilgang og fotfeste i ei lang rekkje ukrainske verksemder ved hjelp av ei skadevare distribuert som del ei oppdatering av ei skatterapporteringsprogramvare. Dataormen NotPetya vart aktivert på alle systema der åtakarane hadde denne tilgangen. Etter at ho var aktivert på ei datamaskin, kartla skadevara nettverket maskina var kopla til, og freista å spreie seg til så mange andre maskiner i nettverket som mogleg. NotPetya henta fyrst ut innloggingsdetaljar (brukarnamn og passord) frå minnet til maskina og freista å nytte desse (på fire ulike måtar) til å få tilgang til andre maskiner over nettverket. Dersom denne framgangsmåten ikkje fungerte, forsøkte skadevara å

[17] Robert S. Mueller, III, *Report On The Investigation Into Russian Interference In The 2016 Presidential Election,* Bd. I (Washington, D.C.: U.S. Department of Justice, 2019), 38–41; United States v. Netyksho Indictment, United States District Court for the District of Columbia, Indictment No. 1:18-cr-215, Doc. 1 (13. juli 2018), avsn. 24–31.

få tilgang ved å utnytte to sikkerheitshol i Windows som på det tidspunktet var relativt nyoppdaga og difor relativt vanlege (kalla EternalBlue og EternalRomance). Samtidig krypterte (og i praksis øydela) NotPetya ein essensiell del av harddisken og sette maskina til å gjere omstart i laupet av 10–60 minutt. Då maskina restarta, var ho i praksis ubrukeleg. Det vil seie at NotPetya etter aktiveringa utførte stega *sidevegs forflytting* og *handlingar på målet* autonomt.[18]

Med utgangspunkt i at målet med operasjonen var å gjere størst mogleg skade, illustrerer NotPetya ein av fordelane med autonom skadevare; spreiing og hastigheit som ikkje er mogleg dersom menneskeleg involvering er naudsynt. På grunn av den øydeleggjande effekten vart skadevara beinveges oppdaga. Ho vart raskt sampla og analysert, mottiltak vart raskt identifisert, og åtaket var over etter eit par dagar. For åtakaren ville det naturleg nok ikkje vore mogleg å råke eit like stort volum på like kort tid dersom åtaket vart utført med manuelle metodar; for eit ikkje-diskriminerande, destruktivt åtak som dette gjer autonomien at skadepotensialet blir mykje større. Det er også mogleg å sjå føre seg at det i ein situasjon med knappheit på personell vil kunne vere nyttig med cybervåpen som er *fire and forget*, altså som ein ikkje treng å nytte merksemd på etter at dei er «avfyrte».

Desse eigenskapane har også nokre openberre baksider. Når åtakaren «slepp laus» ormen, har han samtidig gjeve frå seg kontrollen. Det gjer at skaden råkar vilkårleg, og at spreiinga er uføreseieleg. For NotPetya kan denne vilkårlegheita illustrerast ved at sjølv om NotPetya var eit russisk åtak på Ukraina, spreia skadevara seg langt utover Ukrainas grenser og råka også russisk industri og næringsliv. Dersom ein ønskjer å kunne kontrollere effekten, ser denne typen ormar ut til å vere ein dårleg taktikk.[19]

Eit eksempel på ein dataorm som vart utvikla med motsett hensikt, det vil seie eit målretta åtak, er skadevara som går under namnet Stuxnet. Stuxnet er truleg den mest avanserte autonome skadevara verda kjenner til, og også ein av dei skadevarene som er grundigast studert. Stuxnet var del av den amerikansk-israelske operasjonen Olympic Games, som hadde sabotasje mot det iranske uranopprikingsanlegget i Natanz som mål. Skadevara var

[18] Tekniske detaljar om NotPetya finst i Igal Gofman, «Advanced Threat Analytics security research network technical analysis: NotPetya», *Microsoft Security*, 03.10.2017, https://www.microsoft.com/security/blog/2017/10/03/advanced-threat-analytics-security-research-network-technical-analysis-notpetya/ og Microsoft Defender Security Research Team, «New ransomware, old techniques: Petya adds worm capabilities», *Microsoft Security*, 27.06.2017, https://www.microsoft.com/security/blog/2017/06/27/new-ransomware-old-techniques-petya-adds-worm-capabilities/; den større historia om NotPetya og operasjonen kan lesast om i Andy Greenberg, *Sandworm: A New Area of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (New York: Doubleday, 2019), 179–226.

[19] Greenberg, *Sandworm*, 183; sjå også Conti og Raymond, *On Cyber*, 220–224.

spesifikt designa for å operere autonomt i dei isolerte datanettverka i anlegget; tilgangen var truleg via infiserte minnepinnar eller berbare datamaskiner som personell medvitent eller umedvitent tok med seg inn i fjellanlegget. Ved etablering gjorde Stuxnet sjekkar av operativsystem og anna programvare på målet for å velje beste veg vidare. Skadevara hadde seks ulike metodar for sidevegs forflytting: fire for forflytting internt på eit datanettverk, ein for forflytting (med minnepinne) mellom maskiner på ulike nettverk og ein spesielt designa for å forflytte seg frå det administrative datanettverket til programmerbare logiske styringar (PLS-ar) som styrte uranopprikingsprosessen. Stuxnet manipulerte programmeringa på PLS-ane slik at dei fekk to nye funksjonar. Skadevarekoden som PLS-ane vart infisert med, monitorerte fyrst for ein bestemt operasjonsmodus og gjorde opptak av normale verdiar for prosessen. Så, etter 13 eller 27 dagar, spann han i ein kort periode sentrifugane i opprikingsanlegget med destruktiv fart (alternerte mellom høg og låg fart) samtidig som verdiane frå normaloperasjon vart formidla til kontrollfunksjonen til anlegget, før han gjekk inn i ein ny monitorieringsperiode.[20]

På same måte som NotPetya var Stuxnet ein orm som spreidde seg og sette i verk dei destruktive handlingane sine autonomt. Men der NotPetya spreidde seg og øydela udiskriminert, var Stuxnet laga spesifikt for uranopprikingsanlegget i Natanz. Ormen gjorde detaljerte sjekkar av konfigurasjonane til målsystemet, og han utførte dei destruktive handlingane berre dersom sjekkane tyda på at han var på riktig system.

Den store fordelen med autonomien til Stuxnet var at skadevara hadde evne til å utføre rekognosering, sidevegs forflytting og dei destruktive handlingane sine på dei isolerte datanettverka og -systema utan ein kommando- og kontrollkanal. Dersom ei skadevare skal fjernstyrast av ein operatør, må målsystemet i praksis vere kopla til internett, noko som ofte ikkje er tilfelle for system med stor militær verdi. I tillegg vil ein kommando- og kontrollkanal alltid utgjere ein risiko for operasjonen. Kommunikasjon mellom skadevare og operatør kan avsløre skadevara, og den kan nyttast til å spore operasjonen tilbake til operatøren. Vidare kan kommando- og kontrollkanalen avskjerast eller manipulerast slik at operatøren heilt eller delvis mister kontroll over skadevara. Denne risikoen kan reduserast ved

[20] For ein teknisk gjennomgang av Stuxnet, sjå Nicolas Falliere, Liam O. Murchu og Eric Chien, *W32.Stuxnet Dossier*, Version 1.3 (Cupertino: Symantec, 2010); Historia om Operation Olympic Games er fortalt i David E. Sanger, *Confront and Conceal: Obama's Secret Wars and Surprising Use of American Power* (New York: Crown, 2012), 188–225 og Kim Zetter, *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon* (New York: Crown, 2014).

å minimere kommunikasjonen mellom skadevara og operatøren, og ved at skadevara ikkje er avhengig av kommandoar frå operatøren for å få jobben gjort.[21]

Operation Olympic Games klarte over tid å forstyrre og forsinke det iranske uranopprikingsprogrammet og blir i hovudsak rekna for å vere ein vellykka operasjon. Det er likevel mogleg å peike på utfordringar med konseptet med autonome, målretta dataormar, illustrert av Stuxnet og Olympic Games. Trass innebygde avgrensingsmekanismar spreidde også Stuxnet seg utover det som var målet for operasjonen: Rundt 100 000 infiserte datamaskiner i ei rekkje land har vorte registrert. Sjølv om infeksjonane utanfor Natanz ikkje utførte skade på nokre system, bidrog dei til at skadevara vart oppdaga, og til at operasjonen vart kompromittert. Utfordringa ligg i at sjølv om Stuxnet var målretta, var den også avhengig av spreiing, og fullstendig kontroll over spreiinga er vanskeleg utan menneskeleg kontroll.[22]

Prisen for å lage ei så spesialisert og målretta autonom skadevare er behov for ekstremt detaljert etterretning om målsystemet og grundig utprøving. Utviklinga av Stuxnet gjekk over fleire år og inkluderte testar på ein replika av anlegget i Natanz, sett opp med same typar sentrifugar. Utfordringa kan også illustrerast med at Stuxnet vart laga i fleire versjonar og hadde ein innebygd mekanisme for distribusjon av oppdateringar. Denne mekanismen var basert på både ein kommando- og kontrollkanal over internett (truleg til berbare datamaskiner teke inn og ut av anlegget) og distribusjon mellom maskiner på lokalnettverket internt på anlegget. Stuxnet var difor ikkje fullstendig autonom: Kapabilitetane til skadevara dekkjer ikkje heile prosessen i Figur 1, og ho var underlagt indirekte menneskeleg kontroll.[23]

Kunstig intelligens i offensive cyberoperasjonar

Sjølv om det vil vere mogleg å seie at ei skadevare som Stuxnet viser ein grad av intelligent oppførsel, vil det vere ein regelbasert intelligens og ein refleksjon av intelligente utviklarar. I dag assosierer ein gjerne intelligente eller smarte system med maskinlæringsmetodar, og då spesielt med *djup læring* (maskinlæringsmetodar som nyttar store kunstige nevrale nettverk). Det blir sagt mykje om *potensialet* for å utnytte djup læring og andre maskinlæringsteknikkar i offensive cyberoperasjonar, men litteraturgjennomgangar viser få konkrete eksempel. Det eksisterer einskilde konseptprov som demonstrerer bruk av maskinlæring på einskildsteg i prosessen, slik som smarte, men avgrensa søk etter sikkerheitshol, gjetting av passord og

[21] Conti og Raymond, *On Cyber*, 236–239.
[22] Falliere, Murchu og Chien, *W32.Stuxnet Dossier*, 5–7.
[23] Falliere, Murchu og Chien, *W32.Stuxnet Dossier*, 17–19, 21–22, 47–49; Zetter, *Countdown to Zero Day*, 320–322, 359–360.

generering av åtaksvektorar (særleg phishing). Likevel er desse døma langt frå kravet til autonomi karakterisert over, nemleg taktiske avgjerder basert på observasjonar (sjølv om ideen, særleg om intelligent skadevare som autonomt lærer om målsystemet og tek avgjerder om åtaksvektorar og sidevegs forflytting, er til stades).[24] Maskinlæringsmetodar er avhengige av treningsdata for å gje gode resultat. Sjølv om eit gjeve datasystem eller ein gjeve teknologi i ein bestemt konfigurasjon vil kunne vere ei lukka verd, vil variasjonen i moglege konfigurasjonar og oppsett av datanettverk og -system med ulike protokollar og programvarer for ulike føremål truleg vere så stor at det essensielt er ei open verd. Det verkar difor lite sannsynleg at ein skal kunne utvikle eit datasett og trene ei skadevare til autonomt å kartleggje eit datanettverk ho ikkje har kjennskap til, overvinne alle hindringar og mottiltak, skjule seg for dynamiske forsvararar og finne vegen til eit abstrakt definert mål. Det vil òg vere utfordringar med å trene ei slik skadevare på operasjonelle datasystem (få tilstrekkeleg tilgang) eller simulerte system (skape tilstrekkeleg realisme).[25]

Defensive cyberoperasjonar

På same måte som for offensive cyberoperasjonar kan ein skildre ein overordna modell for defensive cyberoperasjonar. Som nemnt er det ikkje naudsynleg så stor skilnad på sivile og militære organisasjonar, så denne modellen kan i stor grad basere seg på sivil beste praksis. National Institute of Standards and Technology (NIST) har publisert den tonegjevande standarden *Computer Security Incident Handling Guide*, som skildrar ein overordna prosess for handtering av cybersikkerheitshendingar.[26] Adnan mfl. har kartlagt dei typiske arbeidsoppgåvene til operatørar i cybersikkerheitssenter,[27] medan Trent mfl. har undersøkt

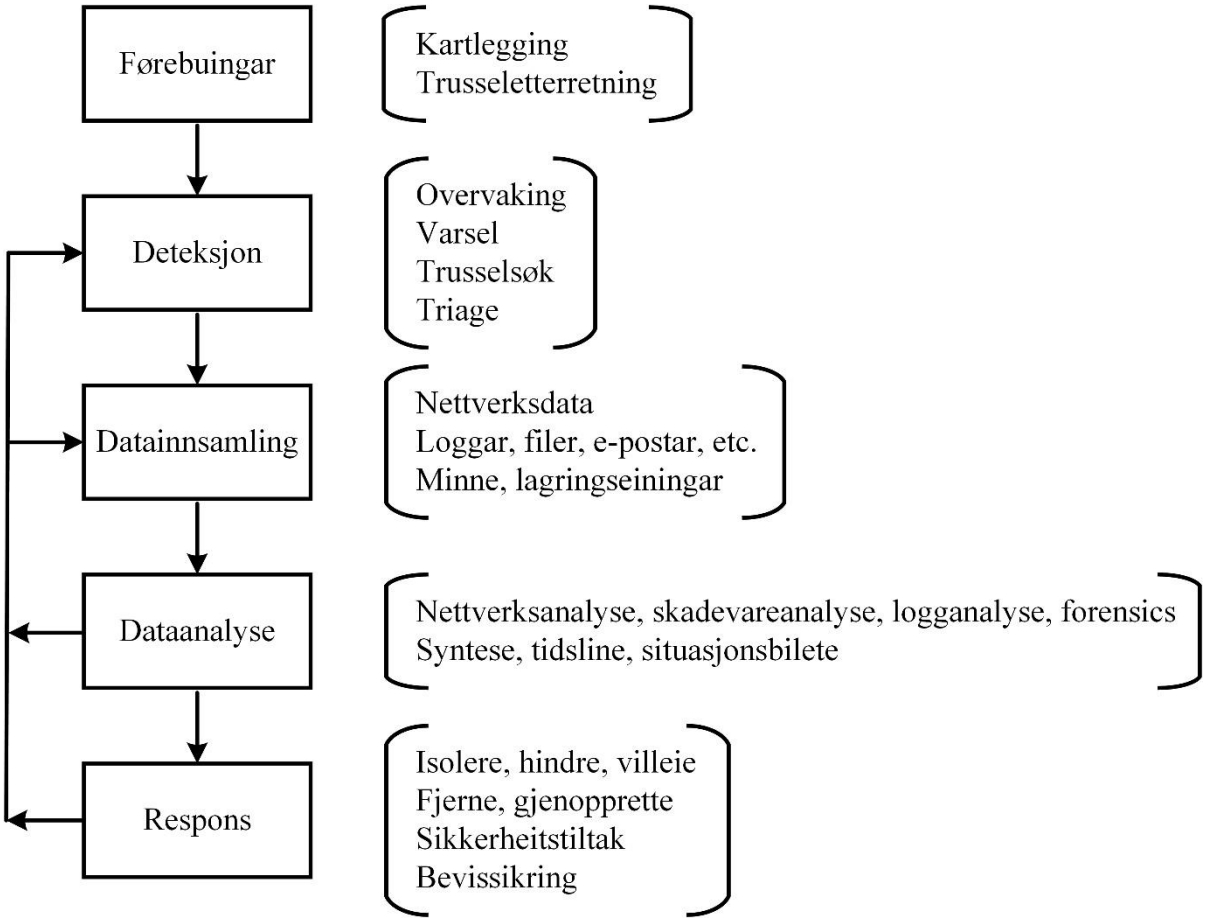
[24] Ben Buchanan mfl., *Automating Cyber Attacks: Hype and Reality* (Washington, D.C.: Center for Security and Emerging Technology, 2020), 11–20, https://doi.org/10.51593/2020CA002; Nektaria Kaloudi og Jingyue Li, «The AI-Based Cyber Threat Landscape: A Survey», *ACM Computing Surveys* 53, nr. 1 (2021), artikkel 20, 7–8, 21–24, https://doi.org/10.1145/3372823; Thanh Cong Truong og Ivan Zelinka, «A Survey on Artificial Intelligence in Malware as Next-Generation Threats», *MENDEL Soft Computing Journal* 25, nr. 2 (2019): 29–30, https://doi.org/10.13164/mendel.2019.2.027; Muhammad Mudassar Yamin mfl., «Weaponized AI for cyber attacks», *Journal of Information Security and Applications* 57 (2021), artikkel 102722, 7–8, https://doi.org/10.1016/j.jisa.2020.102722

[25] Buchanan mfl., *Automating Cyber Attacks,* 23–24.

[26] Paul Cichonski mfl., *Computer Security Incident Handling Guide*, Special Publication 800-61, Rev. utg. 2 (Gaithersburg: National Institute of Standards and Technology, 2012), 21, http://doi.org/10.6028/NIST.SP.800-61r2

[27] Muhammad Adnan mfl., «Investigating the work practices of network security professionals», *Information & Computer Security* 23, nr. 3 (2015): 354–356, https://doi.org/10.1108/ICS-07-2014-0049

arbeidsprosessane i militære defensive cyberoperasjonar.[28] Basert på desse kjeldene kan ein modell for defensive cyberoperasjonar sjå ut som vist i Figur 2.



Figur 2 Defensive cyberoperasjonar

Denne modellen viser den overordna og noko forenkla arbeidsprosessen til eit responsteam[29] – ein dedikert organisasjon eller gruppe personar med ansvar for å forhindre, detektere og handtere cybersikkerheitshendingar – saman med eksempel på dei vanlegaste aktivitetane ein kan finne i dei ulike stega. På same måte som for offensive cyberoperasjonar består kvart steg i denne prosessen av operasjonar utført av menneskelege operatørar støtta av ulike typar programvare, og kvart steg involverer avgjerder om påfølgjande steg. Det kan vere verdt å merkje seg at sjølv om målet med *respons* ofte vil vere å gjenopprette normalfunksjon av datasystema som er under åtak, kan dette steget skjule større taktiske avgjerder. I Forsvarets

[28] Stoney Trent mfl., «Modelling the Cognitive Work of Cyber Protection Teams», *Cyber Defense Review* 4, nr. 1 (2019): 129–131, https://doi.org/10.1108/ICS-07-2014-0049

[29] På engelsk nyttast omgrep som *Incident Response Team* (IRT), *Computer Emergency Response Team* (CERT) og *Cyber Security Incident Response Team* (CSIRT).

Konsept for defensive cyberoperasjonar er *gjenopprette* éi av seks definerte responsar saman med *avskjere*, *sinke*, *isolere*, *hindre* og *nøytralisere*.[30]

Deteksjon vil ofte involvere alarmar frå nettverksensorar eller observasjonar av unormal oppførsel hjå datasystem eller brukarar, men i båe tilfelle er det relativt stort sannsyn for utslag på avvik som ikkje er reelle åtak, med andre ord falske alarmar. Deteksjon vil difor inkludere *triage* av alarmar og observasjonar for å avgjere om ein skal eskalere hendinga til vidare *datainnsamling* og -*analyse*. For eksempel vart det i 2018 teke ei avgjerd om å utplassere nettverkssensorar i datasystema til fylkesmannembeta (nå statsforvaltarane) kort tid etter at ein hadde *detektert* eit pågåande datainnbrot, for å leggje til rette for *datainnsamling* og -*analyse*. Tilsvarande vil analyse av hendinga vere grunnlaget for avgjerder som kva som er beste *respons*. I handteringa av hendinga hjå fylkesmannembeta vart det eit reelt dilemma om ein skulle freiste å fjerne den uautoriserte aktøren frå systema så raskt som mogleg, eller om ein skulle overvake aktøren og gjere meir datainnsamling og analyse før ein avsløra for han at han var oppdaga. Valet falt på det siste, og systema vart overvaka ein månads tid – til ein såg ny aktivitet frå aktøren og valde å kutte internettilgangen for å hindre aktøren i å eksfiltrere data.[31]

Ein endå meir eksplisitt illustrasjon av taktiske vurderingar i ein defensiv cyberoperasjon kan vi finne i den dramatiske forteljinga om Operasjon Bivrost. Ei stor verksemd som handterer sensitive data fekk bistand frå sikkerheitsavdelinga i Telenor Norge til å fjerne ein trusselaktør «med tilnærmet ubegrensede ressurser» frå verksemda sine datasystem. Aktøren hadde fått operere uforstyrra over lang tid og hadde etablert fotfeste. Tiltaka mot aktøren og aktørens svar på dei «er nesten som å spille sjakk».[32]

Sjølv etter at responsteamet har stengt aktøren ute frå det mest sensitive systemet, er teamet medvitne om at jobben enno ikkje er gjort:

> *Vi er sikre på at aktørene vil komme tilbake og forsøke å omgå tiltakene som nå beskytter hans primærmål. På et tidspunkt klarer han akkurat det, men han blir oppdaget gjennom synligheten vi har etablert med sensorer. Dermed klarer vi det på*

[30] Forsvaret, *Konsept for defensive cyberoperasjoner* (Lillehammer: Cyberforsvarets våpenskole, 2022), 20.
[31] Janita A. Bruvoll, Aasmund Thuv og Geir Enemo, *Håndtering av IKT-sikkerhetshendelsene i Helse Sør-Øst og fylkesmannsembetene*, FFI-rapport 20/01560 (Kjeller: Forsvarets forskningsinstitutt, 2020), 37–39, 54–56.
[32] Nikita [pseud.], «Operasjon Bivrost», i *Digital Sikkerhet 2020: de lange linjene* (Telenor, 2020), 48–50, https://www.telenor.no/binaries/om/digital-sikkerhet/Telenor_Digital_Sikkerhet_2020_1.pdf; Telenor; «Slik gikk Telenor Norges sikkerhetsteam til aksjon mot statlige aktører», 30.12.2021, https://www.telenor.no/bedrift/aktuelt/sikkerhet/operasjon-bivrost/

ny; vi utbedrer tiltakene og stikker kjepper i hjulene for ham nok en gang ved å stenge for ham.[33]

Etter andre forsøk er responsteamet framleis ikkje kvitt aktøren og «Operasjon Bivrost besluttes […] vi har bekjempet ham to ganger, nå skal han ut – for godt». Etter omfattande oppgraderingar av systema og tre månader med førebuing og planlegging blir operasjonen iverksett. Alle aktørens tilgangar blir blokkerte, og all skadevare blir fjerna frå systema. I tillegg blir alle passord resette, og ytterlegare sikkerheitsmekanismar blir aktivert. Det er «[e]n klassisk utkastelsesoperasjon», og den må «skje raskt og fortrinnsvis på en tid av døgnet når trusselaktøren normalt ikke jobber. Vi gjør det en helg i november».[34]

Eksempel på autonome defensive cyberoperasjonar

Sjølv om støtteverktøya til operatørane kan ha monaleg med automatikk, vil det fyrst vere når programvare gjer den typen avgjerder involvert i prosessen over, at operasjonane blir å rekne for autonome. Til dømes vil ein nettverkssensor automatisk søkje gjennom større mengder nettverksdata, men det er fyrst når programvare bestemmer kva sensoren skal sjå etter, eller korleis ein alarm frå sensoren skal handterast, at det blir autonomi.
Sikkerheitsorganisering, -automatisering og -respons (SOAR) er eit konsept som viser til integrasjon av ulike støtteverktøy nytta i defensive cyberoperasjonar. Fleire SOAR-system som kan hjelpe eit responsteam med å automatiserer arbeidsflyten i større eller mindre grad, er kommersielt tilgjengeleg, og SOAR utgjer såleis eit potensial for autonomi.[35]

Automated investigation and response (AIR)-funksjonaliteten som Microsoft har i sine Defender-produkt, er eit eksempel på SOAR som har ein grad av autonomi. Automatiske varsel om til dømes policy-brot eller anomaliar vil starte automatisk datainnsamling og -analyse av involverte dataobjekt. Basert på analysen kan AIR autonomt setje i verk tiltak som å blokkere program, filer, e-post, web-linkar og nettverkstilgangar eller sperre brukarkontoar. Under Russlands invasjon av Ukraina i 2022 engasjerte Microsoft seg for å støtte ukrainske defensive cyberoperasjonar. Som eit eksempel detekterte og blokkerte Microsoft Defender i mars 2022 ei tidlegare ukjent destruktiv skadevare som russiske hemmelege tenester freista å levere til eit ukrainsk reiarlag i Lviv. Slik funksjonalitet har

[33] Nikita [pseud.], «Operasjon Bivrost», 48.
[34] Nikita [pseud.], «Operasjon Bivrost», 49–50.
[35] Johnson Kinyua og Lawrence Awuah, «AI/ML in Security Orchestration, Automation and Response: Future Research Directions», *Intelligent Automation & Soft Computing* 28, nr. 2 (2021): 528, https://doi.org/10.32604/iasc.2021.016240

nokre openberre fordelar. Microsoft legg vekt på at det frigjer operatørane i responsteamet frå enkle oppgåver slik at dei kan konsentrere seg om alvorlege hendingar, at responsen blir raskare enn om ein menneskeleg operatør skal gjere datainnsamling og -analyse og setje i verk tiltak, og at Defender, som i eksempelet, kan stoppe åtak i ein tidleg fase og såleis hindre åtakaren i å etablere seg og utføre handlingar i systema.[36]

Mellom baksidene må ein rekne kostnadane ved falske alarmar. Dersom eit autonomt forvarsystem som dette er dårleg konfigurert og ikkje tilpassa verksemda det skal verne, kan falske alarmar som blokkerer program og filer eller sperrar ute brukarar, potensielt gå ut over produktiviteten. I operasjonelle verksemder vil slike falske alarmar kunne bli eit problem for tidskritiske funksjonar og oppgåver.

Eit eksempel på at funksjonalitet som AIR ikkje naudsynleg er tilstrekkeleg for å avverje eller handtere cyberåtak, kan ein finne i det som vert kalla LOLC-hendinga i Norfund. Norfund – eit norsk statleg investeringsfond som investerer i næringsverksemd i utviklingsland – vart i 2020 lurt til å overføre cirka 100 millionar kroner til ein ukjent aktør i staden for den kambodsjanske mikrofinansinstitusjonen LOLC, som var rette mottakar av pengane. Eit sentralt element i svindelen var at svindlarane hadde fått tilgang til e-postkontoen til ein tilsett i Norfund og hadde aktivert vidaresendingsfunksjonen, slik at dei fekk vidaresendt medarbeidarens e-post til eigen e-postkonto. På denne måten kunne dei overvake e-postkommunikasjonen mellom Norfund og LOLC. E-postsystemet til Norfund nytta Microsoft Exhange og var sett opp med automatisk varsling av mistenksam e-postaktivitet. Tre slike varsel vart generert frå den kompromitterte e-postkontoen, og til slutt vart e-postkontoen også automatisk stengt for utgåande e-post. Årsakene til varsla vart ikkje undersøkt i nokre av tilfella, og e-postkontoen vart opna igjen av dei med ansvar for IKT-drift. Det at varsla ikkje vart eskalert til vidare analyse, og at stenging av e-postkontoen vart handsame som produksjonshemmande støy, var medverkande til at kompromitteringa av e-postkontoen ikkje vart avdekka før svindelen vart avslørt over ein månad seinare.[37]

[36] Eric Horvitz, «Applications for Artificial Intelligence in Department of Defense Cyber Missions», *Microsoft On the Issues*, 03.05.2022, https://blogs.microsoft.com/on-the-issues/2022/05/03/artificial-intelligence-department-of-defense-cyber-missions/; Daniel Simpson mfl., «Automated investigation and response in Microsoft 365 Defender», *Microsoft Learn*, 27.09.2022, https://docs.microsoft.com/en-us/microsoft-365/security/defender/m365d-autoir

[37] PwC, *Independent Assessment of the LOLC Incident: Report Developed for Norfund*, Rapport (Oslo: PwC, 2020), 13, 24; Riksrevisjonen, *Riksrevisjonens undersøkelse av informasjonssikkerhet i Norfund,* Rapportvedlegg til Dokument 3:2 (2020–2021) (Oslo: Riksrevisjonen, 2020), 4, 10.

Kunstig intelligens i defensive cyberoperasjonar

Maskinlæringsteknikkar har dei siste åra fått stor merksemd i defensive cyberoperasjonar, i takt med den generelle optimismen knytt til kunstig intelligens i dataintensive domene. Så langt har applikasjonsområda vore konsentrert om einskildoperasjonar i modellen vist ovanfor, og då særleg for aggregering av trusseletterretning, for overvaking (identifikasjon av indikatorar på cyberåtak i nettverkstrafikk og systemoppførsel) og for deteksjon av skadevare.[38] AIR-funksjonaliteten i Microsoft Defender er eit eksempel på slik bruk av kunstig intelligens. I eksempelet frå reiarlaget i Lviv var det maskinlæringsmodellar som tillét Defender å detektere ei skadevare som ikkje tidlegare var registrert, ved å kjenne att teikna på skadevare i data frå systemet.[39] Den store fordelen til Microsoft ved utvikling og trening av kunstig intelligens for defensive cyberoperasjonar – og kanskje også føresetnaden – er den store utbreiinga av programvara deira, og dei enorme mengdene tekniske data dei kan hente ut. Problemet er redusert til deira eiga plattform – som dei har full kontroll over. Samtidig er ho ikkje kva plattform som helst, og gode mekanismar for å avdekke cyberåtak mot Microsoft-produkt vil ha mykje å seie for mange defensive cyberoperasjonar.[40]

Utover den forma for blokkeringar som Microsoft Defender utfører, er det derimot lite progresjon i bruk av kunstig intelligens for å styre prosessane involvert i defensive cyberoperasjonar; kommersielle SOAR-system baserer seg på definerte og regelstyrte arbeidsflytar der avgjerder er gjevne på førehand eller overlatne til den menneskelege operatøren.[41] Bresniker mfl. peikar på at lite har blitt gjort for å fange operatøranes arbeidsprosessar på ein måte som gjer dei eigna for maskinlæring. Vidare argumenterer dei for at måla for defensive cyberoperasjonar ofte er upresise og samansette, og at det difor er utfordrande å formalisere dei på ein måte som maskinlæringsalgoritmar kan nyttiggjere seg av. Resultatet er at det ikkje eksisterer gode treningsdata for å kunne trene ein kunstig

[38] Sjå t.d. Kinyua og Awuah, «AL/ML in Security Orchestration, Automation and Response», 534–539; Thanh Cong Truong mfl., «Artificial Intelligence and Cybersecurity: Past, Presence, and Future», i *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, red. Subhransu Sekhar Dash mfl. (Singapore: Springer, 2020), 354–358, https://doi.org/10.1007/978-981-15-0199-9_30; Zhimin Zhang mfl., «Artificial intelligence in cyber security: research advances, challenges, and opportunities», *Artificial Intelligence Review* 55 (2022): 1035–1037, https://doi.org/10.1007/s10462-021-09976-0

[39] Horvitz, «Applications for Artificial Intelligence»; Ruofan Wang og Kelly Kang, «AI-driven adaptive protection against human-operated ransomware», *Microsoft Security*, 15.11.2021, https://www.microsoft.com/en-us/security/blog/2021/11/15/ai-driven-adaptive-protection-against-human-operated-ransomware/

[40] Arie Agranonik, Shay Kels og Guy Arazi, «Seeing the big picture: Deep learning-based fusion of behavior signals for threat detection», *Microsoft Security*, 23.07.2020, https://www.microsoft.com/en-us/security/blog/2020/07/23/seeing-the-big-picture-deep-learning-based-fusion-of-behavior-signals-for-threat-detection/

[41] Kinyua og Awuah, «AI/ML in Security Orchestration, Automation and Response», 535, 539.

intelligens til å utføre arbeidsprosessane i defensive cyberoperasjonar.[42] Sommer og Paxson legg vekt på at det eksisterer eit semantisk gap mellom indikatorar og anomaliar som kunstige intelligensar kan avdekke, og den *tolkinga* av desse som ein operatør må gjere for å forstå det som eit cyberåtak, og at dette gapet er ei avgjerande og uløyst utfordring.[43]

Dette er i tråd med konklusjonane til Zhong mfl., som har eksperimentert med autonom *triage* av hendingar. Dei samla sekvensar av analyseoperasjonane som operatørar utfører på nettverktrafikk i ein triage, og nytta desse til å konstruere generelle algoritmar for autonom triage. Konklusjonen frå eksperimenta var at sjølv dette innleiande steget av ein defensiv cyberoperasjon er for komplekst og komplisert til å gjerast fullt ut autonomt. Den grunnleggjande utfordringa er at variasjonen i dei mønstera ein menneskeleg operatør vil sjå etter, er stor og må generaliserast frå store mengder data frå tidlegare åtak, samtidig som spesifikk kunnskap om det forsvarte systemet er naudsynt. Gapet mellom tekniske data og abstrakte åtaksmetoder er det framleis den menneskelege operatøren som må fylle.[44] Sjølv om dette arbeidet ikkje var basert på maskinlæring og kanskje kan reknast for å vere ein slags regelbasert intelligens, støytte dei på eit tilsvarande gap som det skildra over, mellom kva ein får maskinene til å gjere autonomt, og dei analysane og fortolkingane av dataa som er naudsynt i defensive cyberoperasjonar.

Vegen fram til eit system som autonomt og ved hjelp av kunstig intelligens kan utføre dei taktiske stega i defensive cyberoperasjonar, som til dømes hendingshandteringa i fylkesmansembeta eller Operasjon Bivrost, synest difor lang.

Avslutning

I cyberdomenet – som er ei «verd av datasystem» der programvarestøtte og automatikk er ein grunnleggande bestanddel – er det ikkje naudsynleg openbert kva ein skal rekne for autonomi. Modellar og eksempel på offensive og defensive cyberoperasjonar viser at cyberoperasjonar kan ha ein grad av autonomi, der ein forstår autonomi som det å ta avgjerder og utføre handlingar for å kome nærmare operasjonens mål, utan direkte instruksjonar frå ein menneskeleg operatør.

---

[42] Kirk Bresniker mfl., «Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cybersecurity», *Computer* 52, nr. 12 (2019), 46–50, https://doi.org/10.1109/MC.2019.2942584

[43] Robin Sommer og Vern Paxson, «Outside the Closed World: On Using Machine Learning for Network Intrusion Detection», i *2010 IEEE Symposium on Security and Privacy* (IEEE Computer Society, 2010), 308, https://doi.org/10.1109/SP.2010.25

[44] Chen Zhong, John Yen og Peng Liu, «Can Cyber Operations Be Made Autonomous? An Answer from the Situational Awareness Viewpoint» i *Adaptive Autonomous Secure Cyber Systems*, red. Sushil Jajodia mfl. (Cham: Springer, 2020), 85–86, https://doi.org/10.1007/978-3-030-33432-1_4

Samtidig er det avgrensingar i denne autonomien. Cyberoperasjonar føreset detaljert kjennskap og god forståing av datasystema som skal angripast eller forsvarast, og denne kjennskapen og forståinga kan ikkje programvare sjølv opparbeide. Vidare kan autonom programvare ta avgjerder om kva som er gode nestesteg innanfor ein overordna taktikk viss omgjevnaden og rammene er kjente, men ho er ikkje i stand til sjølv å velje eller utarbeide taktikk. Taktikken må uansett kvile på menneskelege operatørar, og omsetjinga frå dei måla ein organisasjon set seg for ein operasjon, til mål som autonom programvare kan handtere, er det framleis menneske som må gjere.

Det er stor interesse for bruk av kunstig intelligens og maskinlæring i cyberoperasjonar, inkludert idear om at maskinlæring skal «løfte» operasjonane til ei «høgare» form for autonomi som også handterer overordna taktiske avgjerder, men så langt har utnyttinga av slike metodar vore avgrensa til automatisering av einskildsteg innanfor ein taktikk.

Gjennomgangen har vist at spørsmålet om autonomi ikkje er anten eller. Det gjev meining å snakke om gradar av autonomi, og ein kan forvente at cyberoperasjonar gradvis utnyttar potensialet som ligg i automatisering av operatøranes oppgåver. Ein definisjon av autonomi i cyberoperasjonar vil vere kontekstavhengig, og med atterhald om at kapittelet er skrive med samtidas blikk, er cyberoperasjonar i sin natur semiautonome, med gradar av automatikk og menneskeleg kontroll. Automatisering og autonomi gjev stort potensial for effektivisering og skalering av cyberoperasjonar, for handtering av store datamengder og for frigjering av arbeidskapasitet hjå operatørane, men i lang tid vil menneskelege operatørar vere avgjerande for den etterretninga og dei taktiske vala ein vellykka cyberoperasjon er avhengig av.[45]

Litteratur

Adnan, Muhammad, Mike Just, Lynne Baillie og Hilmi Gunes Kayacik. «Investigating the work practices of network security professionals». *Information & Computer Security* 23, nr. 3 (2015): 347–367. https://doi.org/10.1108/ICS-07-2014-0049

Agranonik, Arie, Shay Kels og Guy Arazi. «Seeing the big picture: Deep learning-based fusion of behavior signals for threat detection». *Microsoft Security*. 23.07.2020. https://www.microsoft.com/en-us/security/blog/2020/07/23/seeing-the-big-picture-deep-learning-based-fusion-of-behavior-signals-for-threat-detection/

[45] Buchanan mfl., *Automating Cyber Attacks*, 22–23; Conti og Raymond, *On Cyber*, 217–226.

Berkeley, Edmund C. *Giant Brains or Machines That Think.* New York: John Wiley & Sons, 1949.

Bresniker, Kirk, Ada Gavrilovska, James Holt, Dejan Milojicic og Trung Tran. «Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cybersecurity». *Computer* 52, nr. 12 (2019): 45–52. https://doi.org/10.1109/MC.2019.2942584

Bruvoll, Janita A., Aasmund Thuv og Geir Enemo. *Håndtering av IKT-sikkerhetshendelsene i Helse Sør-Øst og fylkesmannsembetene*. FFI-rapport 20/01560. Kjeller: Forsvarets forskningsinstitutt, 2020.

Buchanan, Ben, John Bansemer, Dakota Cary, Jack Lucas og Micah Musser. *Automating Cyber Attacks: Hype and Reality.* Washington, D.C.: Center for Security and Emerging Technology, 2020. https://doi.org/10.51593/2020CA002

Cichonski, Paul, Tom Millar, Tim Grance og Karen Scarfone. *Computer Security Incident Handling Guide.* Special Publication 800-61, Rev. utg. 2. Gaithersburg: National Institute of Standards and Technology (NIST), 2012. https://doi.org/10.6028/NIST.SP.800-61r2

Conti, Gregory og David Raymond. *On Cyber: Towards an Operational Art for Cyber Conflict.* Kopidion Press, 2017.

Falliere, Nicolas, Liam O. Murchu og Eric Chien. *W32.Stuxnet Dossier*, Version 1.3. Cupertino: Symantec, 2010.

Forsvaret. *Forsvarets fellesoperative doktrine*. Oslo: Forsvarsstaben, 2019.

Forsvaret. *Konsept for defensive cyberoperasjoner.* Lillehammer: Cyberforsvarets våpenskole, 2022.

Gofman, Igal. «Advanced Threat Analytics security research network technical analysis: NotPetya». *Microsoft Security*. 03.10.2017. https://www.microsoft.com/security/blog/2017/10/03/advanced-threat-analytics-security-research-network-technical-analysis-notpetya/

Grant, Tim, Ivan Burke og Renier van Heerden. «Comparing Models of Offensive Cyber Operations». I *Leading Issues in Cyber Warfare & Security Research*, Bd. 2, redigert av Julie Ryan, 35–55. Reading, Storbritannia: Academic Conferences and Publishing International, 2015.

Greenberg, Andy. *Sandworm: A New Area of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. New York: Doubleday, 2019.

Haigh, Thomas, Mark Priestley og Crispin Rope. *ENIAC in Action: Making and Remaking the Modern Computer*. Cambridge, MA: MIT Press, 2016.

Horvitz, Eric. «Applications for artificial intelligence in Department of Defense cyber missions». *Microsoft On the Issues*. 03.05.2022. https://blogs.microsoft.com/on-the-issues/2022/05/03/artificial-intelligence-department-of-defense-cyber-missions/

Hutchins, Eric M., Michael J. Cloppert og Rohan M. Amin. «Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains». I *Proceedings of the 6th International Conference on Information Warfare and Security*, redigert av Leigh Armistead, 113–125. Reading, Storbritannia: Academic Publishing International, 2011.

Jajodia, Sushil, George Cybenko, V. S. Subrahmanian, Vipin Swarup, Cliff Wang og Michael Wellman (red.). *Adaptive Autonomous Secure Cyber Systems*. Cham: Springer, 2020.

Joint Chiefs of Staff. *Cyberspace operations*. Joint Publication 3-12. Washington, D.C.: Joint Chiefs of Staff, 2018. https://irp.fas.org/doddir/dod/jp3_12.pdf

Joyce, Rob. «Disrupting nation state hackers». Video, 34:55. *USENIX Enigma*. 27.01.2016. https://www.usenix.org/conference/enigma2016/conference-program/presentation/joyce

Kaloudi, Nektaria og Jingyue Li. «The AI-Based Cyber Threat Landscape: A Survey». *ACM Computing Surveys* 53, nr. 1, artikkel 20 (2021): 1–34. https://doi.org/10.1145/3372823

Kinyua, Johnson og Lawrence Awuah. «AI/ML in Security Orchestration, Automation and Response: Future Research Directions». *Intelligent Automation & Soft Computing* 28, nr. 2 (2021): 527–545. https://doi.org/10.32604/iasc.2021.016240

Lockheed Martin. «The Cyber Kill Chain». Lese 05.09.2022. https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html

Lund, Mass Soldal. «Cyber som operasjonsdomene». *Norsk militært tidsskrift* 186, nr. 1 (2017): 28–34.

Lund, Mass Soldal. «CND-konsept for taktiske nettverk». I *Ledelse i cyberdomenet*, redigert av Øyvind Jøsok og Benjamin J. Knox, 35–47, Lillehammer: Forsvarets ingeniørhøgskole, 2018.

Lund, Mass Soldal. «Menneskemaskina». *Syn og segn* 126, nr. 1 (2020): 74–81.

Lund, Mass Soldal. «Cybertrusselen og elektronisk krigføring – kva kan vi sjå føre oss i gråsona?» *Norsk militært tidsskrift* 191, nr. 3 (2021): 22–29.

Marcus, Gary og Ernest Davis. *Rebooting AI: Building Artificial Intelligence We Can Trust.* New York: Pantheon Books, 2019.

Microsoft Defender Security Research Team. «New ransomware, old techniques: Petya adds worm capabilities». *Microsoft Security*. 27.06.2017. https://www.microsoft.com/security/blog/2017/06/27/new-ransomware-old-techniques-petya-adds-worm-capabilities/

Mitre. «Att&ck». Lese 03.08.2022. https://attack.mitre.org/

Mueller, Robert S., III. *Report On The Investigation Into Russian Interference In The 2016 Presidential Election*, Bd. I. Washington, D.C.: U.S. Department of Justice, 2019.

Nato. *Allied Joint Doctrine for Cyberspace Operations*. Allied Joint Publication 3.20, Utg. A, Versjon 1. Genève: Nato Standardization Office, 2020.

Nikita [pseud.]. «Operasjon Bivrost». I *Digital Sikkerhet 2020: de lange linjene*, 46–51. Telenor, 2020. https://www.telenor.no/binaries/om/digital-sikkerhet/Telenor_Digital_Sikkerhet_2020_1.pdf

PwC. *Independent Assessment of the LOLC Incident: Report Developed for Norfund*. Rapport. Oslo: PwC, 2020.

Riksrevisjonen. *Riksrevisjonens undersøkelse av informasjonssikkerhet i Norfund.* Rapportvedlegg til Dokument 3:2 (2020–2021). Oslo: Riksrevisjonen, 2020.

Ringstad, Jan Terje. *Mind the gap – An exploratory study of commercial and military computer security incident response teams (CSIRTs) – Are Incident Response (IR) and Computer Security Incident Response Teams (CSIRTs) Forensic Ready in the Information Domain?*. Masteroppgåve. Norges teknisk-naturvitenskapelige universitet (NTNU), 2017.

Sanger, David E. *Confront and Conceal: Obama's Secret Wars and Surprising Use of American Power*. New York: Crown, 2012.

Simpson, Daniel, Denise Vangel, Alex Buck, Siddarth Mandalika, Stephanie Savell, Ashok Lobo og Joe Davies. «Automated investigation and response in Microsoft 365 Defender». *Microsoft Learn*. 27.09.2022. https://docs.microsoft.com/en-us/microsoft-365/security/defender/m365d-autoir

Sommer, Robin og Vern Paxson. «Outside the Closed World: On Using Machine Learning for Network Intrusion Detection». I *2010 IEEE Symposium on Security and Privacy*, 305–316. IEEE Computer Society, 2010. https://doi.org/10.1109/SP.2010.25

Telenor. «Slik gikk Telenor Norges sikkerhetsteam til aksjon mot statlige aktører». 30.12.2021. https://www.telenor.no/bedrift/aktuelt/sikkerhet/operasjon-bivrost/

Trent, Stoney, Robert R. Hoffman, David Merritt og Sarah Smith. «Modelling the Cognitive Work of Cyber Protection Teams». *Cyber Defense Review* 4, nr. 1 (2019): 125–135.

Truong, Thanh Cong og Ivan Zelinka. «A Survey on Artificial Intelligence in Malware as Next-Generation Threats». *MENDEL Soft Computing Journal* 25, nr. 2 (2019): 27–34. https://doi.org/10.13164/mendel.2019.2.027

Truong, Thanh Cong, Ivan Zelinka, Jan Plucar, Marek Čandík og Vladimír Šulc. «Artificial Intelligence and Cybersecurity: Past, Presence, and Future». I *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, redigert av Subhransu Sekhar Dash, C. Lakshmi, Swagatam Das, og Bijaya Ketan Panigrahi, 351–363. Singapore: Springer, 2020. https://doi.org/10.1007/978-981-15-0199-9_30

Wang, Ruofan og Kelly Kang. «AI-driven adaptive protection against human-operated ransomware». *Microsoft Security*. 15.11.2021. https://www.microsoft.com/en-us/security/blog/2021/11/15/ai-driven-adaptive-protection-against-human-operated-ransomware/

Yamin, Muhammad Mudassar, Mohib Ullah, Habib Ullah og Basel Katt. «Weaponized AI for cyber attacks». *Journal of Information Security and Applications* 57 (2021): artikkel 102722, 1–14. https://doi.org/10.1016/j.jisa.2020.102722

Zetter, Kim. *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*. New York: Crown, 2014.

Zhang, Zhimin, Huansheng Ning, Feifei Shi, Fadi Farha, Yang Xu, Jiabo Xu, Fan Zhang og Kim-Kwang Raymond Choo. «Artificial intelligence in cyber security: research advances, challenges, and opportunities». *Artificial Intelligence Review* 55 (2022): 1029–1053. https://doi.org/10.1007/s10462-021-09976-0

Zhong, Chen, John Yen og Peng Liu. «Can Cyber Operations Be Made Autonomous? An Answer from the Situational Awareness Viewpoint». I *Adaptive Autonomous Secure Cyber Systems*, redigert av Sushil Jajodia, George Cybenko, V.S. Subrahmanian, Vipin Swarup, Cliff Wang og Michael Wellman, 63–88. Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-33432-1_4