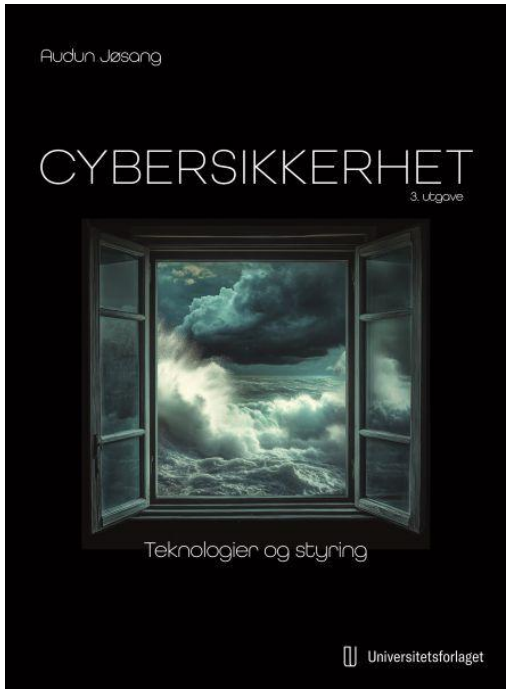
IN1030
Systemer, krav og konsekvenser

Krav om cybersikkerhet



Audun Jøsang
Institutt for Informatikk
Universitetet i Oslo
30. april 2026

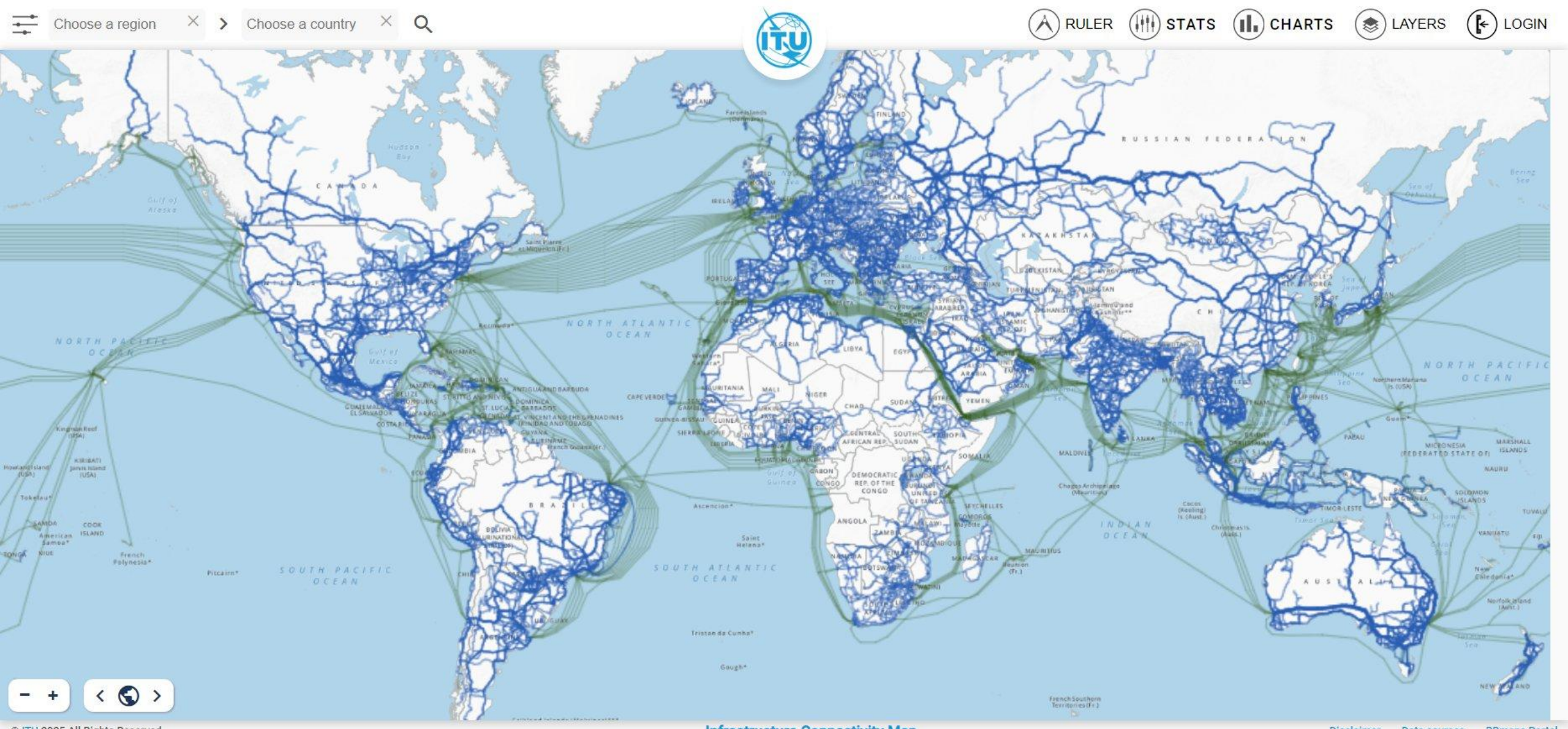
Oversikt

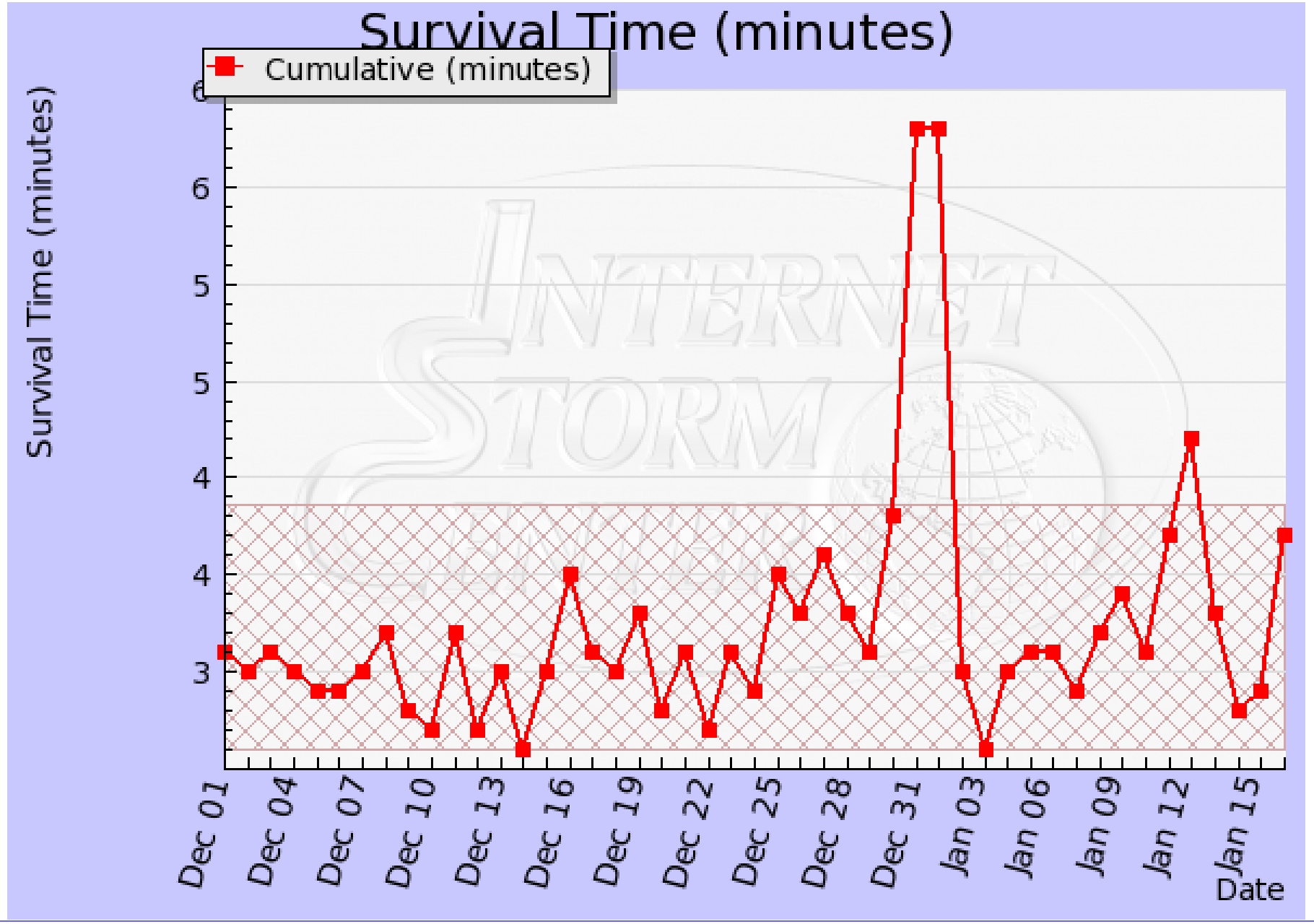
- Internetts grenseløshet og cyberusikkerhet
- Grunnleggende begreper for cybersikkerhet
- Krav om styring av cybersikkerhet
- KI og cyber

Internettets grenseløshet og cyberusikkerhet

Det fysiske internett

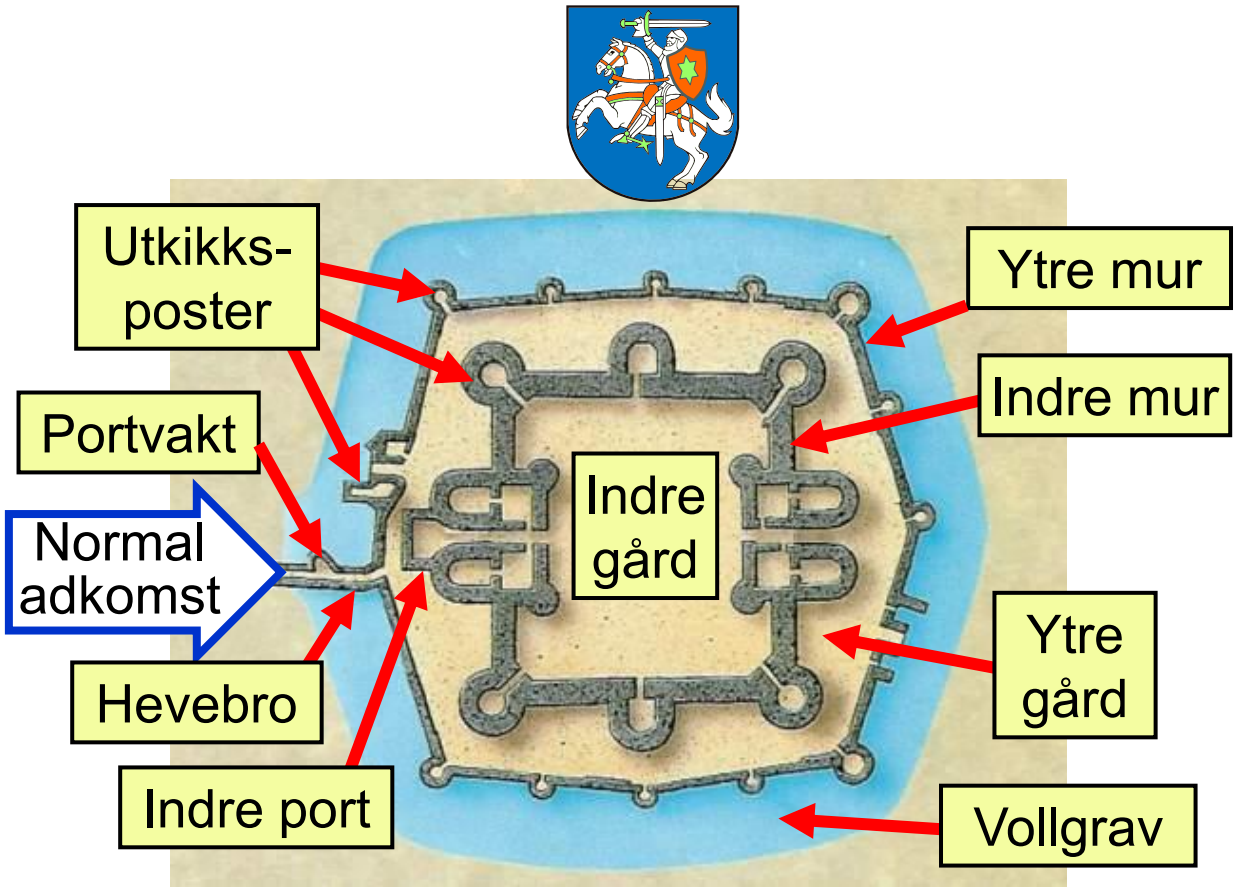
https://bbmaps.itu.int/products



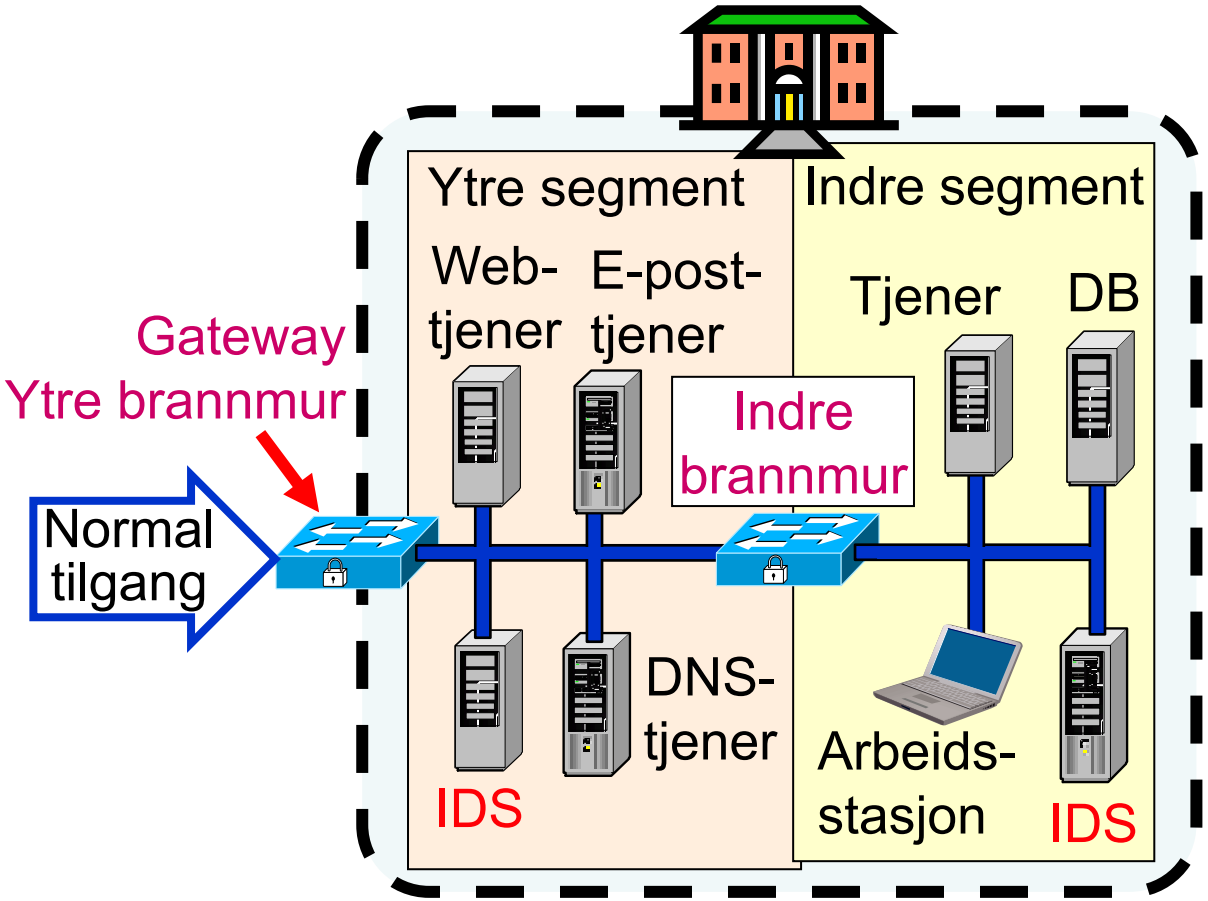


Internet Storm Center: https://www.dshield.org/survivaltime/

Analogi: Middelalderborg og nettverkssikkerhet



Middelalderborg



Datanett

WORLD ECONOMIC FORUM

Reports

Published: 14 January 2026

Global Risks Report 2026

Download PDF

The *Global Risks Report 2026*, the 21st edition of this annual report, marks the second half of a turbulent decade. The report analyses global risks through three timeframes to support decision-makers in balancing current crises and longer-term priorities. Chapter 1 presents the findings of this year's **Global Risks Perception Survey (GRPS)**, which captures insights from over 1,300 experts worldwide. It explores risks in the current or immediate term (in 2026), the short-to-medium term (to 2028) and in the long term (to 2036). Chapter 2 explores the range of implications of these risks and their interconnections, through six in-depth analyses of selected themes. Below are the key findings of the report, in which we compare the risk outlooks across the three-time horizons.

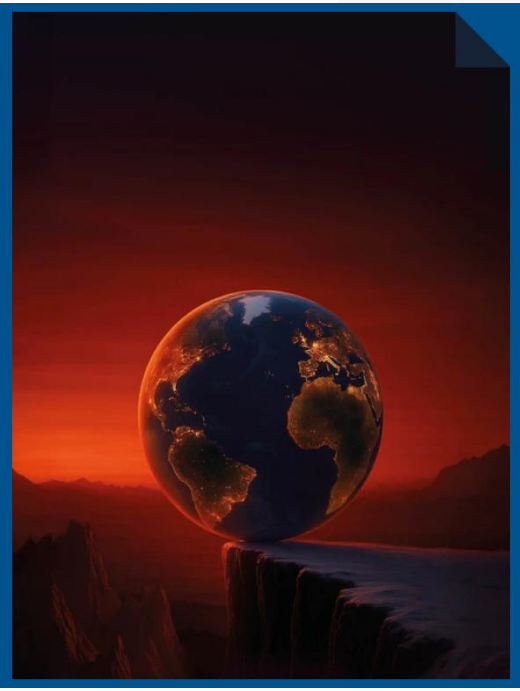


FIGURE 3 **Global risks ranked by severity, short term**

"Please estimate the likely impact (severity) of the following risks

Short term (2 years)

1. Geoeconomic confrontation
2. Misinformation and disinformation
3. Societal polarization
4. Extreme weather events
5. State-based armed conflict
6. Cyber insecurity
7. Inequality
8. Erosion of human rights and/or of civic freedoms
9. Pollution
10. Involuntary migration or displacement

Grunnleggende begreper for cybersikkerhet

Sikkerhet er beskyttelse av verdier mot skade

Eiendom Infrastruktur Demokrati Lov og orden Liv og helse Miljø Digitale verdier Persondata





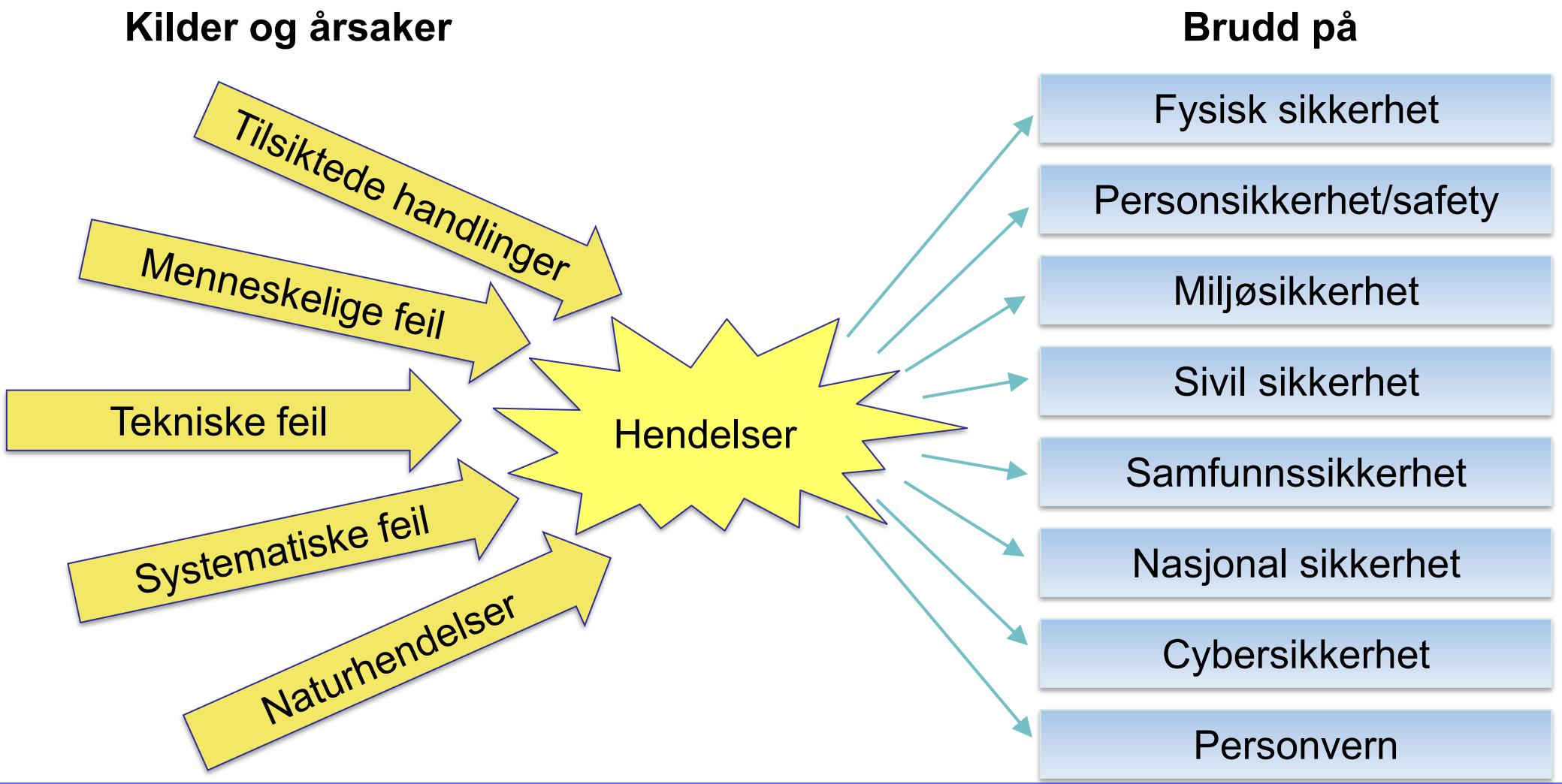





- **Fysisk sikkerhet:** hindre innbrudd, tyveri og tukling med utstyr
- **Samfunnssikkerhet:** opprettholde funksjonalitet i kritiske infrastrukturer
- **Nasjonal sikkerhet:** demokrati, politisk stabilitet, territorial integritet
- **Sivil sikkerhet og rettssikkerhet:** opprettholdelse av lov og orden
- **Personsikkerhet / trygghet / safety:** beskyttelse av liv og helse
- **Miljøsikkerhet:** hindre forurensing og fremmede arter
- **Cybersikkerhet:** beskyttelse av digitale verdier mot skade
- **Personvern:** følge prinsipper for innhenting, lagring, behandling og deling av personopplysninger

Usikkerhet er mangel på beskyttelse av verdier mot skade, som antonym til sikkerhet.

Ulike kilder til brudd på sikkerhet



Hva er *cybersikkerhet* ?



- *Cybersikkerhet* er beskyttelse av *digitale verdier* mot skade. Det inkluderer beredskap for håndtering av hendelser og gjenoppretting når skade skjer.
- Hvilke kategorier av digital verdier skal beskyttes?
  - Cyberforsvarsmatrisen definerer digitale verdier som: ENHETER, APPER, NETTVERK, DATA, og BRUKERE. «Digitale verdier» kalles også «digitale eiendeler».
- Hvordan kan digitale verdier skades?
  - Brudd på ett eller flere av sikkerhetsmålene **Konfidensialitet**, **Integritet** og **Tilgjengelighet** (KIT)
- Dekker både tilsiktet og utilsiktet skade
  - Trusselaktører kan være mennesker eller naturlige hendelser
  - Mennesker kan gjøre skade både tilsiktet og utilsiktet
- Definisjon av informasjonssikkerhet:
  - ***Beskyttelse av informasjonens Konfidensialitet, Integritet og Tilgjengelighet.***

I tillegg kan andre egenskaper, f.eks. autentisitet, sporbarhet, ubenektelighet og pålitelighet omfattes. (ISO/IEC 27000:2018)

Kjært barn har mange navn

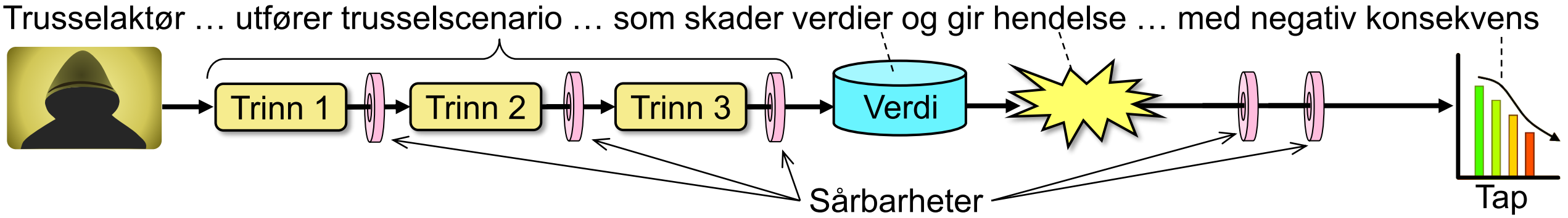


- **Informasjonssikkerhet:** generelt begrep, dekker f.eks. også beskyttelse av informasjon på papir, populært fra 1970-tallet. Direkte oversettelse av «information security» som er nedfelt i en rekke internasjonale standarder og rammeverk.
- **Datasikkerhet:** (eng. Computer Security) kan tolkes som beskyttelse av data, men det betyr egentlig beskyttelse av datamaskiner, populært fra 1980-tallet.
- **IT- og IKT-sikkerhet:** tolkes som sikkerhet i IT/IKT-systemer, populært fra 1990-tallet.
- **Cybersikkerhet:** tolkes ofte som sikkerhet for alt som har å gjøre med Internett og som er koblet til Internett, populært fra 2010-tallet.
- **Digital sikkerhet:** skapt av norske myndigheter i 2019, med hensikt å være et samlebegrep for alle begrepene ovenfor, rimer godt med «digitalisering» og ser ut til å bli populært fremover.

Alle disse begrepene betyr det samme !

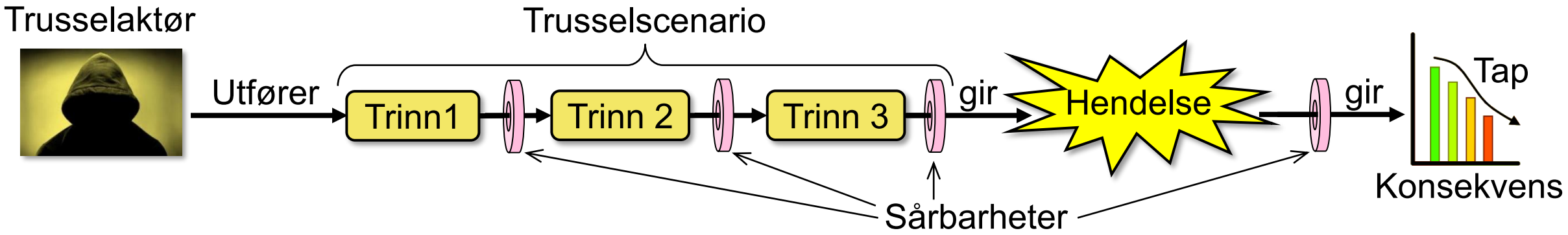
- **Sikring:** I noen sektorer (f.eks. olje og gass) der «sikkerhet» er ensbetydende med «safety» benyttes «sikring» i betydning av «security».

Verdier, Trusler, Sårbarheter og Tiltak

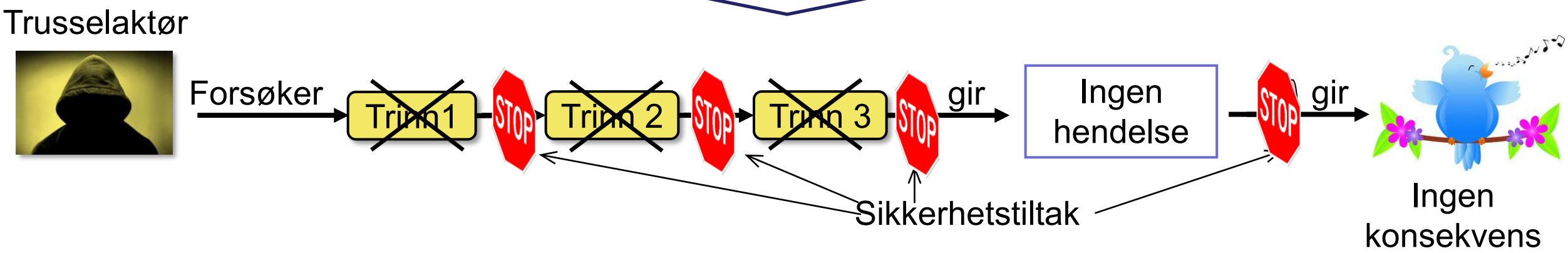


- **Verdier:** Digitale eiendeler som er av verdi for organisasjonen.
  - ENHETER, APPER, NETTVERK, DATA, og BRUKERE
  - Mål for informasjonssikkerhet er å beskytte eiendelenes KIT, avhengig av behov.
  - Person(opplysnings)vern
- **Trussel:** Et potensielt angrepsscenario som kontrolleres av en trusselaktør og som kan skade organisasjonens digitale eiendeler
- **Sårbarhet:** Manglende sikkerhetstiltak mot trusler og evne til å håndtere hendelser.
- **Sikkerhetstiltak** (Security Control): Metode for å forhindre trusler eller redusere konsekvenser av hendelser

Sikkerhetstiltak = fjerne sårbarheter

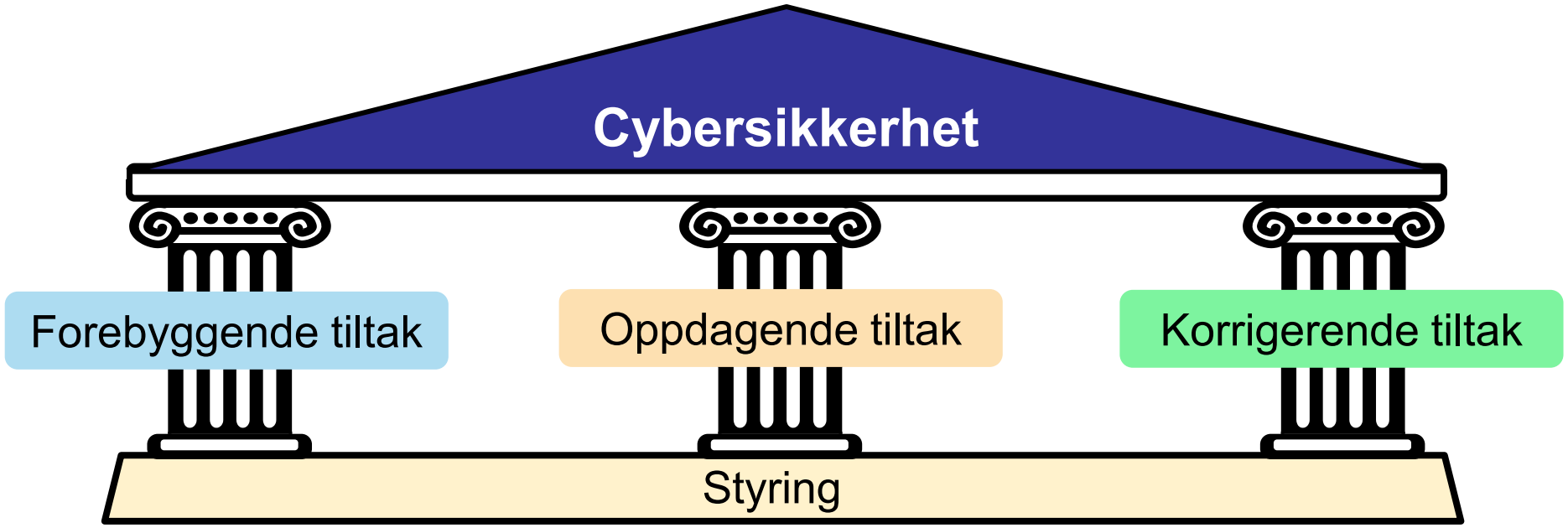


Formål med cybersikkerhetsstyring

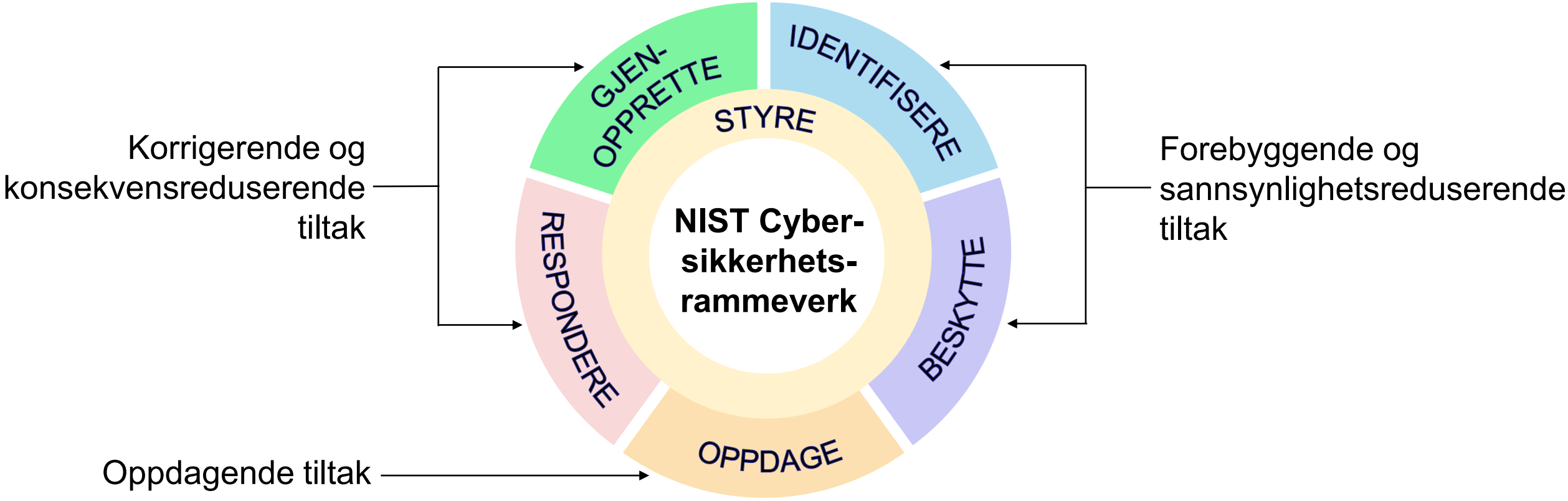


Tiltak for cybersikkerhet

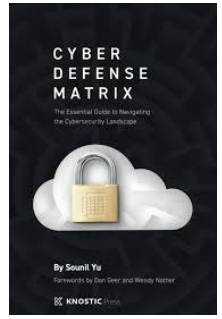
Eksempel kategorisering etter sikkerhetsfunksjoner



NIST Cybersikkerhetsrammeverk
National Institute for Standards and Technology, USA



Cyberforsvarsmatrisen (Cyber Defense Matrix)



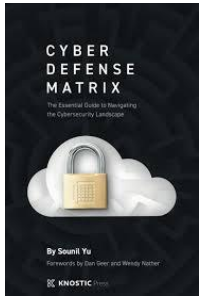
Cybersikkerhetsfunksjoner

Digitale verdier	IDENTIFISERE	BESKYTTE	OPPDAGE	RESPONDERE	GJEN-OPPRETTE
ENHETER	■ ● ●	■ ■ ● ● ●	■ ● ●	■ ● ● ●	■ ● ●
APPER	●	■ ■ ■ ● ● ● ●	■ ● ●	●	●
NETTVERK	■ ● ● ● ●	■ ■ ■ ■ ● ● ● ●	■ ■ ● ● ●	■ ● ●	■ ■ ● ●
DATA	■ ● ●	■ ■ ■ ● ●	●	■ ●	■ ● ●
BRUKERE	■ ● ● ●	■ ● ●	● ●	●	●

Forklaring: ■ Nåværende profil ● Ønsket profil

Digitale verdier (se cyberforsvarsmatrisen, kap.18)

"Digitale verdier" kalles også "digitale eiendeler"

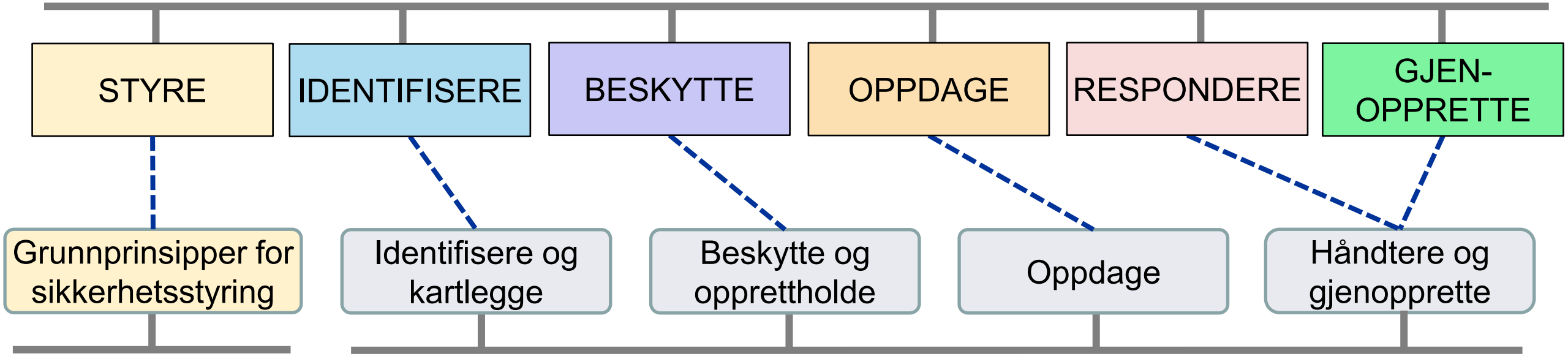


- ENHETER er maskinvare og standard installert programvaren som OS.
  - Fastvare, operativsystemer og standardapplikasjoner fra leverandører
  - Skill mellom ENHETER basert på hvem som eier enheten, for eksempel organisasjon, ansatte (BYOD) eller skyleverandører.
  - Nettverksenheter som svitsjer og rutere tilhører også denne klassen.
- APPER er applikasjoner og prosesser spesielt utviklet av eller for virksomheten.
  - APP-er produserer vanligvis tjenestene og produktene som leverer organisasjonens verdi og forretningsfunksjoner. Derfor er kontinuiteten og integriteten til APPS vanligvis kritisk.
- NETTVERK er kommunikasjonskanaler, rutere, protokoller og brannmurfiltreringer.
  - Konfigurasjon av nettverksenheter som DNS (Domain Name System), AS (Autonomous Systems) og PKI (Public Key Infrastructure), og som gir ruting, filtrering og VPN
- DATA er forretningsdata, som er lagret, som overføres eller brukes av verdiene over.
- BRUKERE er personene og AI-agentene, med ID-er og autentikatorer.
  - En AI-agent kan kanskje betraktes som en enhet, men hvis den har egenskapene til en bruker med sin egen ID, autentiseringslegitimasjon, tilgangsautorisasjoner og intelligente beslutninger.

Rammeverk for cybersikkerhet

National Institute of Standards and Technology
U.S. Department of Commerce

Funksjoner i NISTs Cybersikkerhetsrammeverk



NSM-veileder

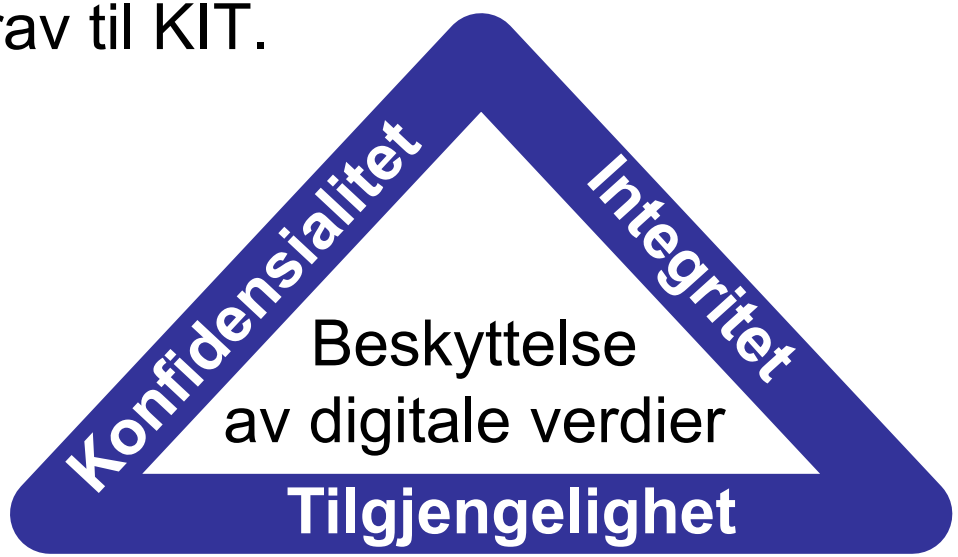
Funksjoner i NSMs grunnprinsipper for IKT-sikkerhet

NSM

Generelle sikkerhetsmål: KIT+P

- Informasjonssikkerhet er tradisjonelt definert som opprettholdelse av KIT.
- Digitale verdier generelt, ikke bare informasjon, har behov for KIT.
- Ulike digitale verdier har ulike krav til KIT.

- Engelsk: CIA
  - **Confidentiality**
  - **Integrity**
  - **Availability:**



- Person(opplysnings)vern (data protection) er et tilleggsmål som bl.a. forutsetter KIT. GDPR (General Data Protection Regulation) definerer krav til personvern.

Personvern

Sikkerhetstmål og sikkerhetstiltak

- Sikkerhetsmål
  - Uavhengig av spesifikk implementering
  - Kan implementers med ulike tiltak/controller
- Sikkerhetstiltak / controller / mekanismer
  - Basert på spesifikk implementering, ofte bundet til spesifikke produkter

Sikkerhetsmål:

Konfidensialitet – Integritet – Tilgjengelighet

støtter

Sikkerhetstiltak:

f.eks. låser – kryptering – autentisering - sikkerhetskultur





Analogi for sivil sikkerhet

Konfidensialitet

- Egenskapen av at informasjon ikke blir gjort tilgjengelig eller vist til uautoriserte individer, entiteter eller prosesser.
  (ISO/IEC 27000)
- Trusler:
  - Datatyveri (ekstern trussel)
  - Datalekkasje (intern trussel).
- Sikkerhetstiltak eksempler:
  - Kryptering,
  - Kryptografiske kommunikasjonsprotokoller, f.eks. TLS
  - Autentisering og tilgangskontroll,
  - Anonymisering, f.eks. gjennom pseudonym eller VPN
  - Skallsikring
  - Sikkerhetskultur, bevissthet
  - …



Integritet

- **Dataintegritet:** Egenskapen av at data ikke har blitt endret eller slettet på en uautorisert måte. (X.800)
- **Systemintegritet:** Egenskapen av å opprettholde korrekthet og kompletthet av dataressurser (ISO/IEC 27000)
- Trusler: Ødelagte data og miskonfigurerte systemer
- Sikkerhetstiltak eksempler:
  - Hashing, MAC, kryptering
  - Konfigurasjonsstyring
  - Endringsledelse
  - Autentisering
  - Tilgangskontroll
  - Sertifisert programvare
  - Sikkerhetskultur, bevissthet
  - …

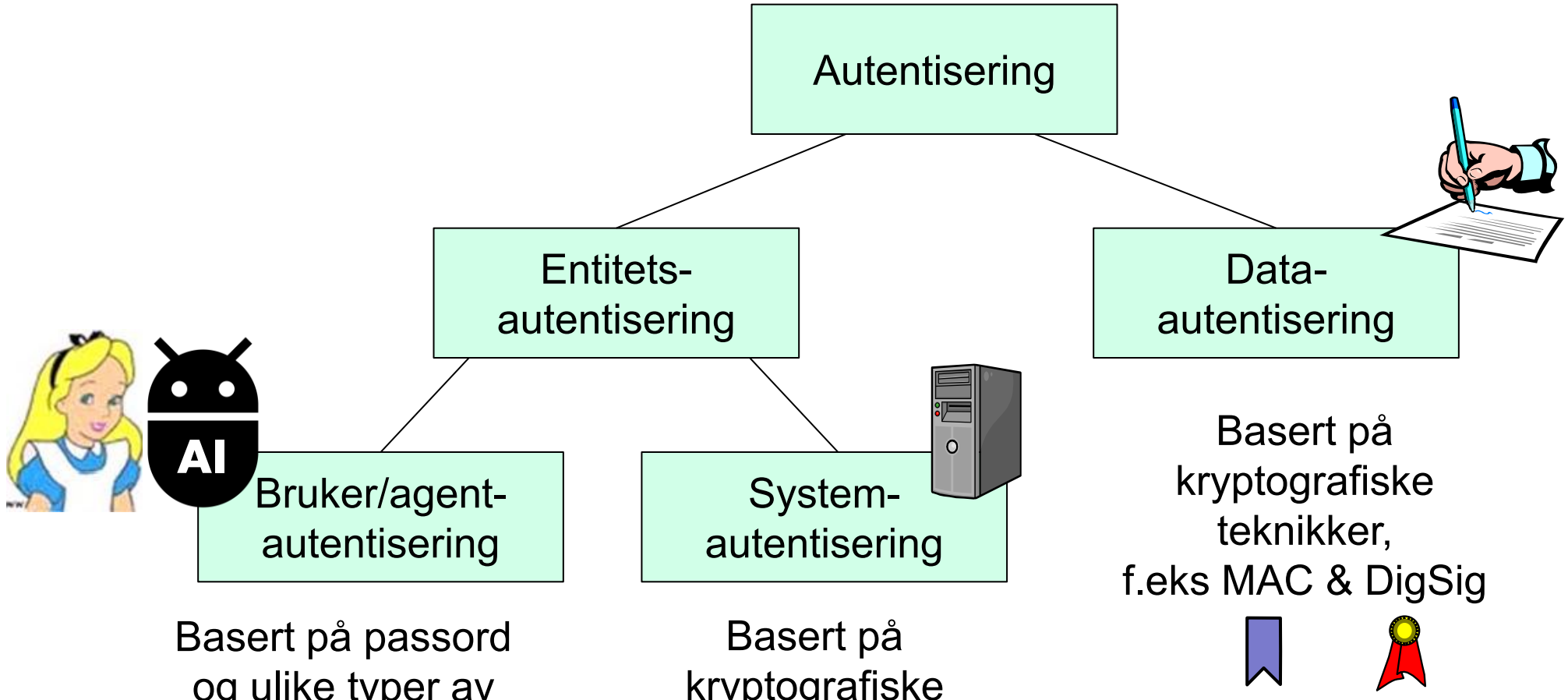


Tilgjengelighet

- Egenskapen av at data og tjenester er tilgjengelige og anvendbare ved forespørsel fra en autorisert entitet. (ISO/IEC 27000)
- Trusler:
  - Tjenestenekt, overlastangrep (DoS / DDoS)
  - Løsepengevirus
  - Forsinkelse av tidskritiske funksjoner.
- Sikkerhetstiltak eksempler:
  - Redundans av ressurser,
  - Failover-konfigurasjon
  - Brannmur
  - Sikkerhetskopiering (backup)
  - Hendelsesrespons og beredskap,



Typer av autentisering



Tilgangsautorisering



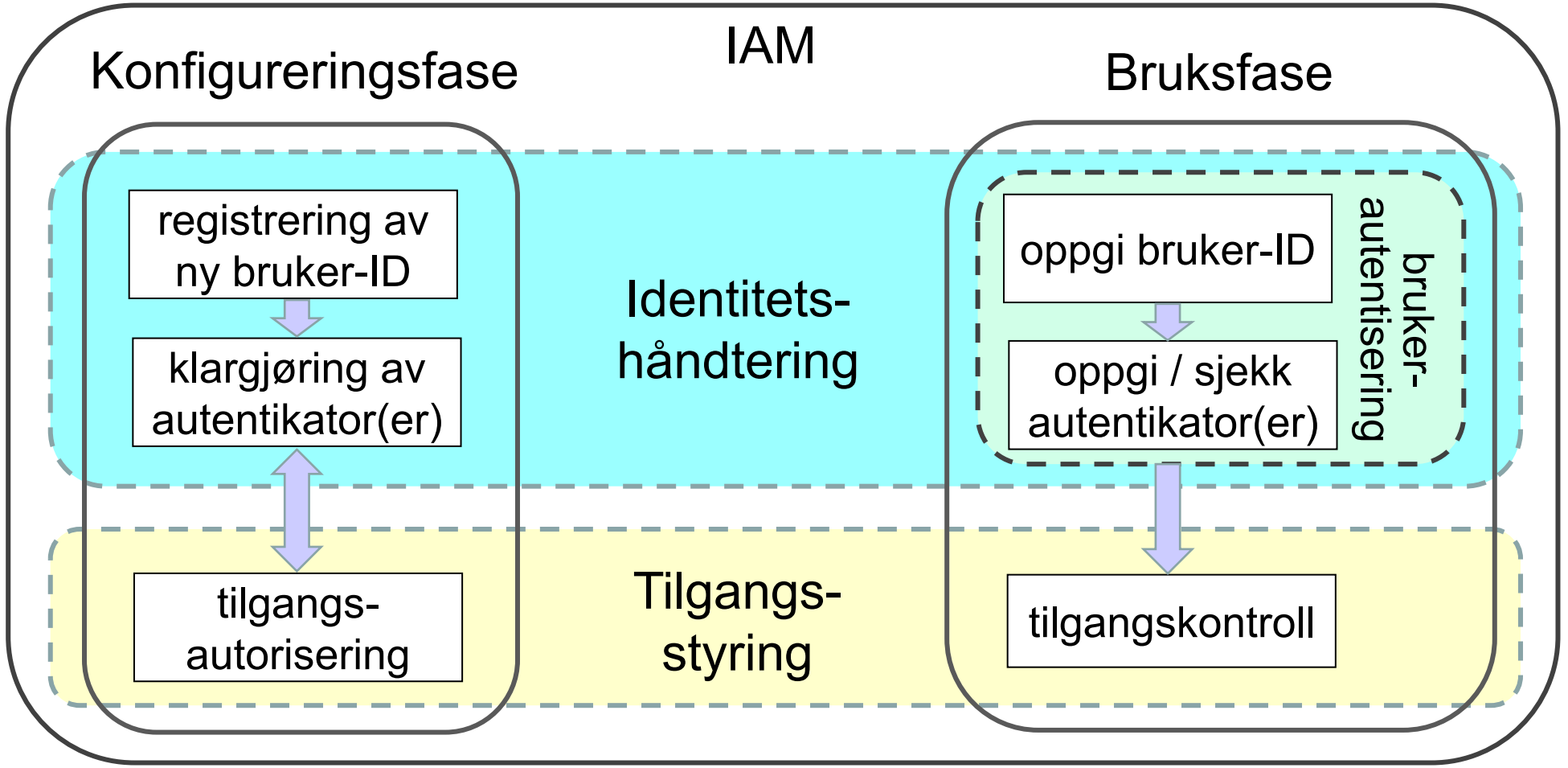
- Tilgangsautorisering er å spesifisere tilgangsrettigheter for entiteter, dvs. for brukere, roller og prosesser
  - Spesifiserer hvem som skal ha tilgang til hva
  - Autoriseringspolicyen er vanligvis definert av mennesker
  - Autoriseringspolicyen blir formalisert som regler og konfigureringer for tilgangskontroll i systemer.
- Autorisering kan bli delegert
  - Leder → Sys.Admin → Bruker
- Vær oppmerksom på forvirring i andre lærebøker og kilder:
  - Noen steder defineres tilgangsautorisering som ekvivalent med tilgangskontroll. Dette er fullstendig feil, fordi den gjør definisjonen på konfidensialitet meningsløs, på den måten at det ikke ville være et brudd på konfidensialitet hvis en hacker får tilgang til en konto med et cracket passord og stjeler data.

Tilgangskontroll

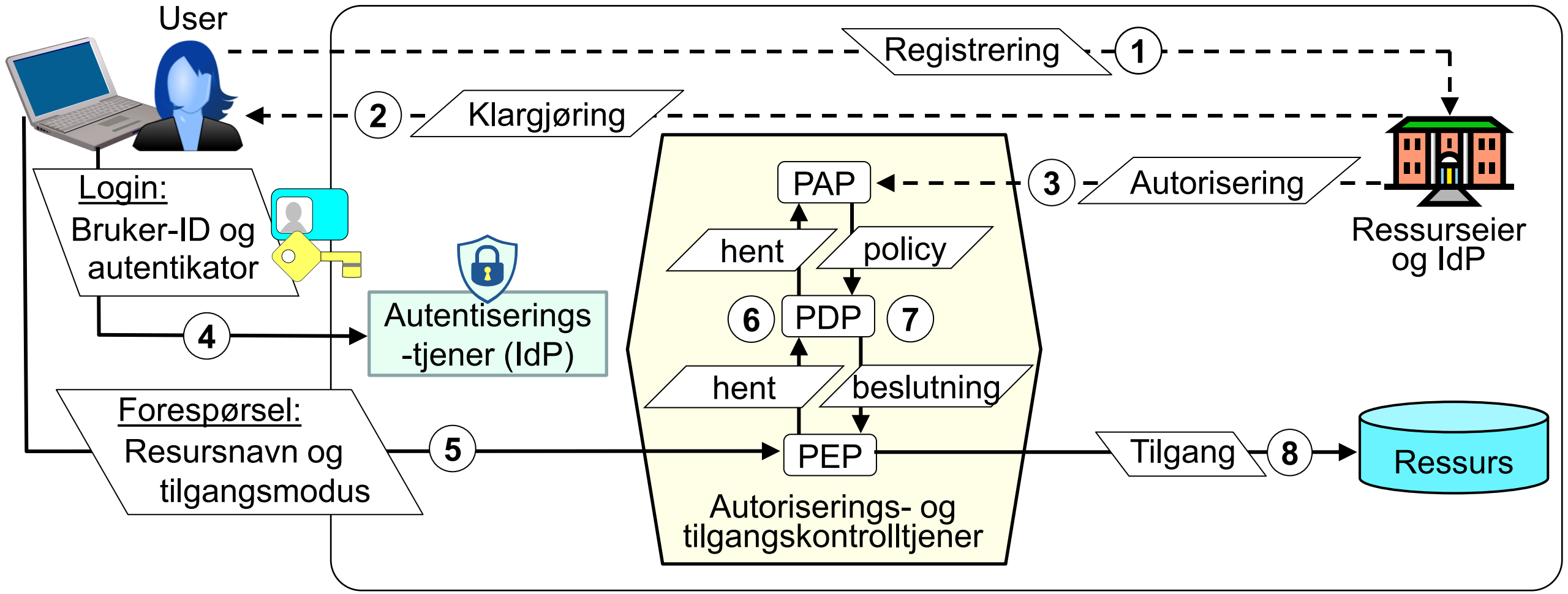


- Tilgangskontroll foregår etter at brukeren er autentisert.
- Brukeren/entiteten må være autentisert for at systemet skal vite hvem som prøver å utføre en handling eller forespør tilgang.
- Tilgangskontroll benytter autoriseringspolicy/regler for å avgjøre om brukeren er autorisert for tilgang til ressurser.
- Policy/regler for tilgangsautorisering defineres under konfigureringsfasen slik at tilgangskontroll kan utføres under bruksfasen.
- Mange ulike måter å definere regler for tilgangskontroll, f.eks.
  - Identitetsbasert (DAC)
  - Merkebasert (MAC)
  - Rollebasert (RBAC)
  - Attributtbasert (ABAC) (generalisering av alle måtene ovenfor)

Identitetshåndtering og tilgangsstyring
IAM (Identity and Access Management)



IAM scenario



PAP: Policy Administration Point
PDP: Policy Decision Point
PEP: Policy Enforcement Point
IdP: Identity Provider
Konfigreringsfase
Bruksfase

Betydning og oversettelse av begrepet “sikkerhet”

Engelsk		Norsk
• Security	→	• Sikkerhet
• Safety	→	• Trygghet
• Certainty	→	• Visshet



Presis oversettelse

- Security
- Safety
- Certainty

→ • Sikkerhet



Upresis oversettelse

Betydning og oversettelse av begrepet “usikkerhet”

Engelsk	Norsk
• Insecurity	• Usikkerhet
• Unsafety	• Utrygghet
• Uncertainty	• Uvisshet

World Economic Forum rangerer cyberusikkerhet som global risiko nr.6



Presis oversettelse

- Insecurity
- Unsafety
- Uncertainty

→ • Usikkerhet



Upresis oversettelse

Krav om styring av cybersikkerhet

Kilder til krav om cybersikkerhet



- **God praksis:** Krav om adekvat sikkerhet i forretningsprosesser i henhold til vanlig god praksis og forvaltning.
  - Vanlig god praksis setter f.eks. krav om brukerautentisering og tilgangskontroll.
  - Veiledere for god praksis er f.eks. NSMs grunnprinsiper, NISTs CSF, og CIS controls (kap.18)
- **Risikovurdering:** Krav om å begrense sikkerhetsrisiko til et forsvarlig nivå.
  - Tiltak identifiseres gjennom risikovurdering, f.eks. krav om 2-faktorautentisering.
  - Pentesting (enkel form for risikovurdering) avdekker sikkerhetssårbarheter som bør fjernes.
- **Regelverk:** Juridiske, lovbestemte, regulatoriske, bransje- og kontraktsmessige krav til informasjonssikkerhet, f.eks.:
  - Sikkerhetsloven setter krav om sikkerhet for grunnleggende nasjonale funksjoner.
  - Lov om digital sikkerhet (NIS2) setter krav om cybersikkerhet for viktige/kritiske funksjoner
  - GDPR setter krav om beskyttelse av persondata.
  - Norm for informasjonssikkerhet og personvern i helse- og omsorgssektoren (bransjenorm).

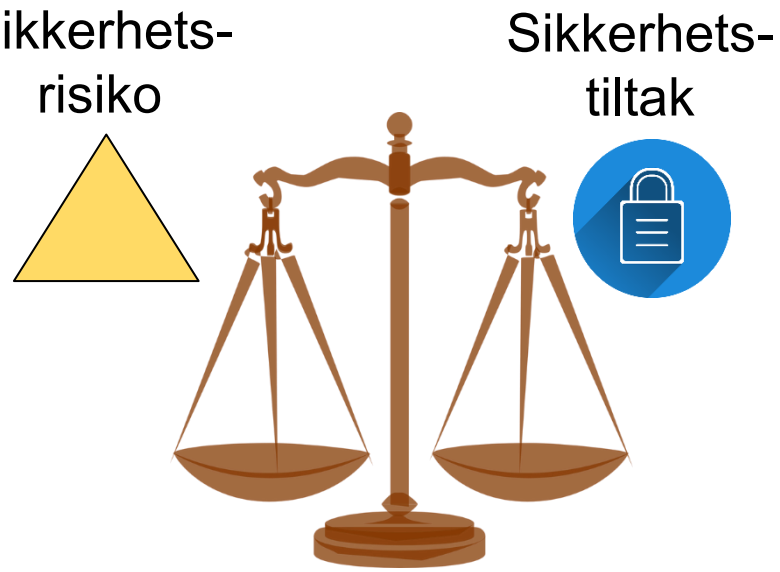




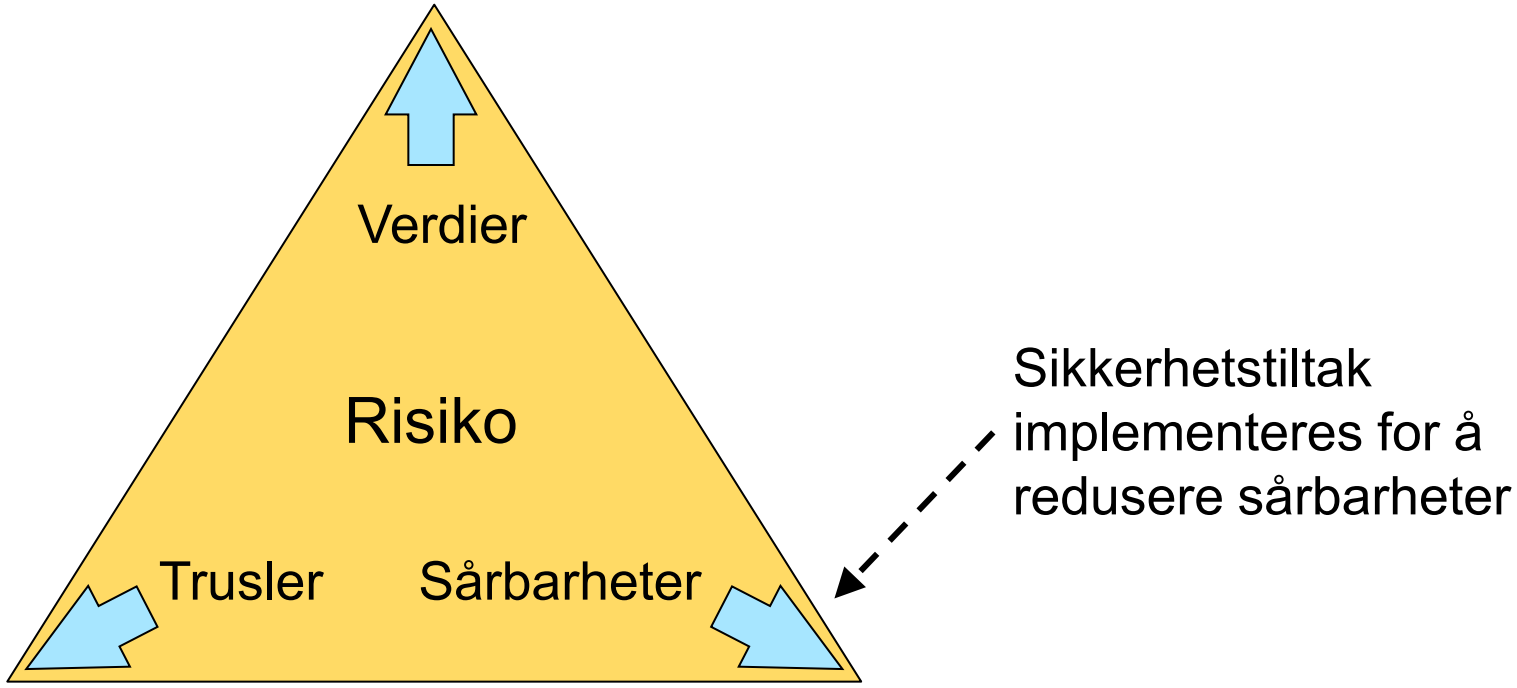
Merk at regelverk og juridiske krav om informasjonssikkerhet ofte henviser til risikovurdering.

Målsetting for styring av cybersikkerhet

- Ville det være mulig å løse alle sikkerhetsproblemer?
- Nei, fordi:
  - Det oppdages stadig nye sårbarheter i gamle systemer
  - Nye digitale tjenester, ofte med sårbarheter, eksponeres online
  - Trusselaktører er flinke til å finne nye sårbarheter som kan utnyttes
  - Det utvikles stadig mer effektive angrepsverktøy
  - Økende antall og alvorlighet av trusler
- Konklusjon: Informasjonssikkerhet er en kontinuerlig prosess for å stoppe trusler og fjerne sårbarheter
- Målsetting for styring av informasjonssikkerhet er å oppnå god balanse mellom sikkerhetsrisiko og sikkerhetstiltak.

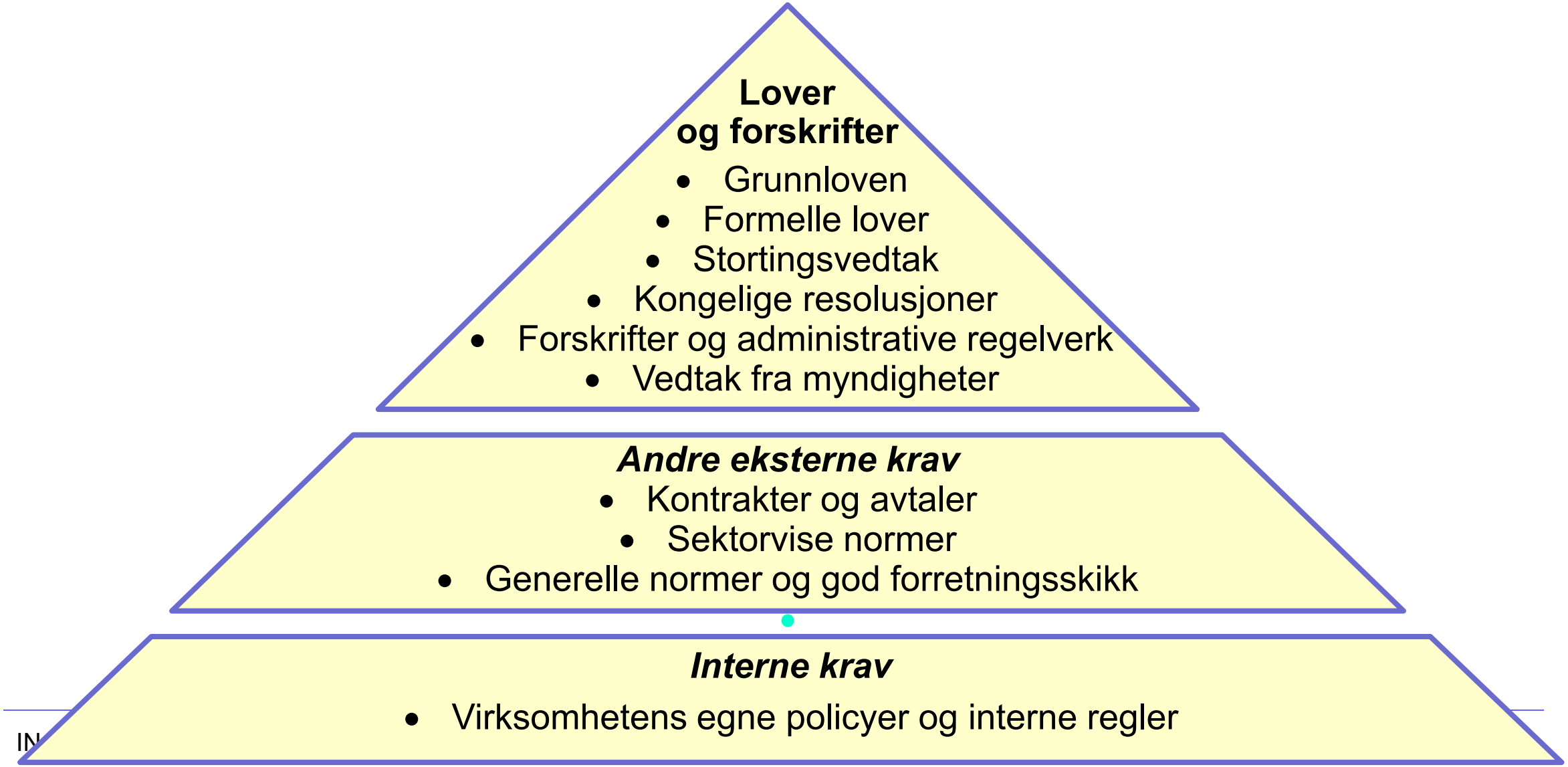


Generell risikomodell for IT-sikkerhet

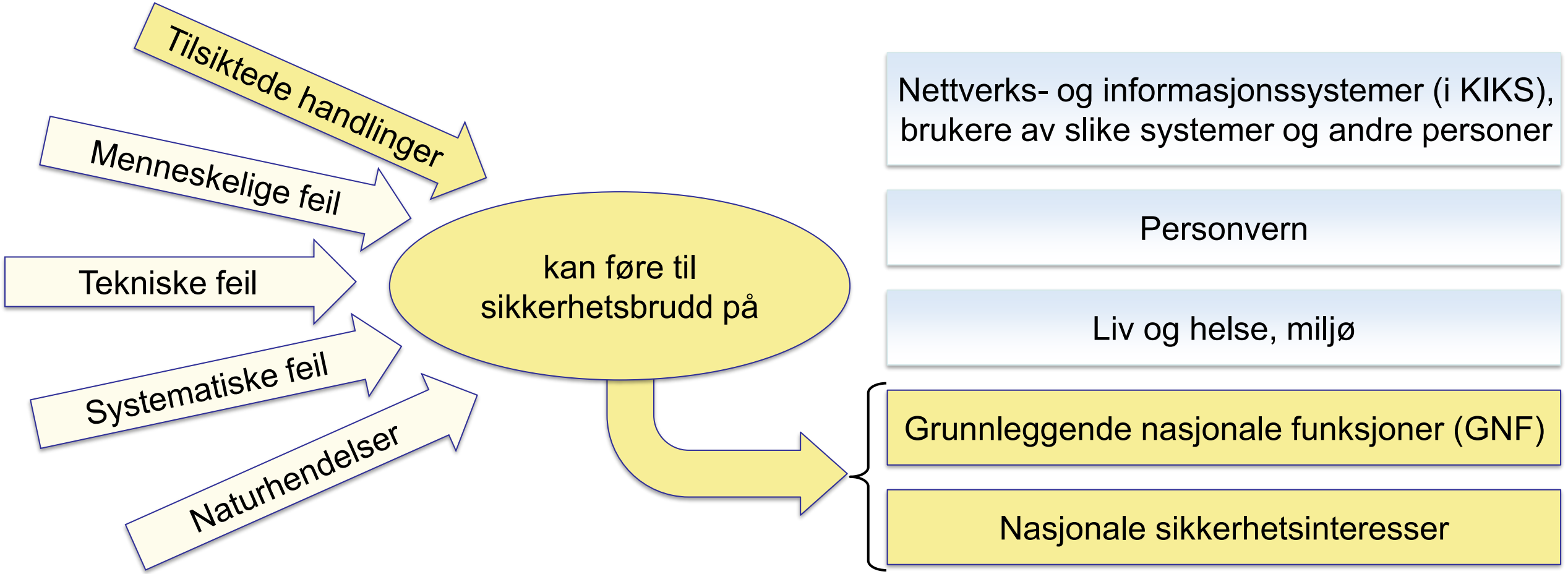


- Generell modell for risiko
  - Jo større og flere verdier du har, jo større og flere trusler du er utsatt for, og jo mere sårbar du er, desto større risikoeksponering har du.

Hierarki av regulatoriske krav



Fokus for sikkerhetsloven (administrert av NSM)



Kravkjede for styring av informasjonssikkerhet

- Sikkerhetsloven, § 4-1.Sikkerhetsstyring
  - Kongen kan gi forskrift om sikkerhetsstyring.
- Virksomhetsikkerhetsforskriften § 3.Styringssystem for sikkerhet
  - En virksomhet som omfattes av sikkerhetsloven, skal etablere et styringssystem for sikkerhet.
- ISO/IEC 27001 Ledelsessystem for informasjonssikkerhet – Krav
  - også kalt styringssystem for informasjonssikkerhet
  - engelsk: ISMS (Information Security Management System)
- Sertifisering av virksomhetens ISMS i henhold til ISO/IEC 27001
  - 7 kategorier med krav
  - Relevanserklæring av ISO/IEC 27002 Informasjonssikkerhetstiltak

Regelverk fra EU



- EU har et omfattende regelverk bestående hovedsaklig av direktiver (eng. directives) og forordninger (eng. acts, regulations).
- Et EU-direktiv skal implementeres i nasjonale lover med omtrent samme innhold.
- Hvis Norge tar inn et nytt direktiv, kan det hende at Norge allerede har lover som dekker direktivet helt eller delvis.
- En EU-forordning skal som hovedregel implementeres som nasjonal lov nøyaktig slik den er artikulert av Europakommisjonen, bare oversatt til det nasjonale språk.
- EØS-landene bestemmer sammen om et EU-direktiv eller en EU-forordning skal gjelde i EØS-landene, som dermed betyr at det blir tatt inn i selve EØS-avtalen.
- Her gjelder prinsippet om alle eller ingen, det vil si at Norge ikke kan bestemme alene om vi skal eller ikke skal ta inn et EU-direktiv eller en EU-forordning.

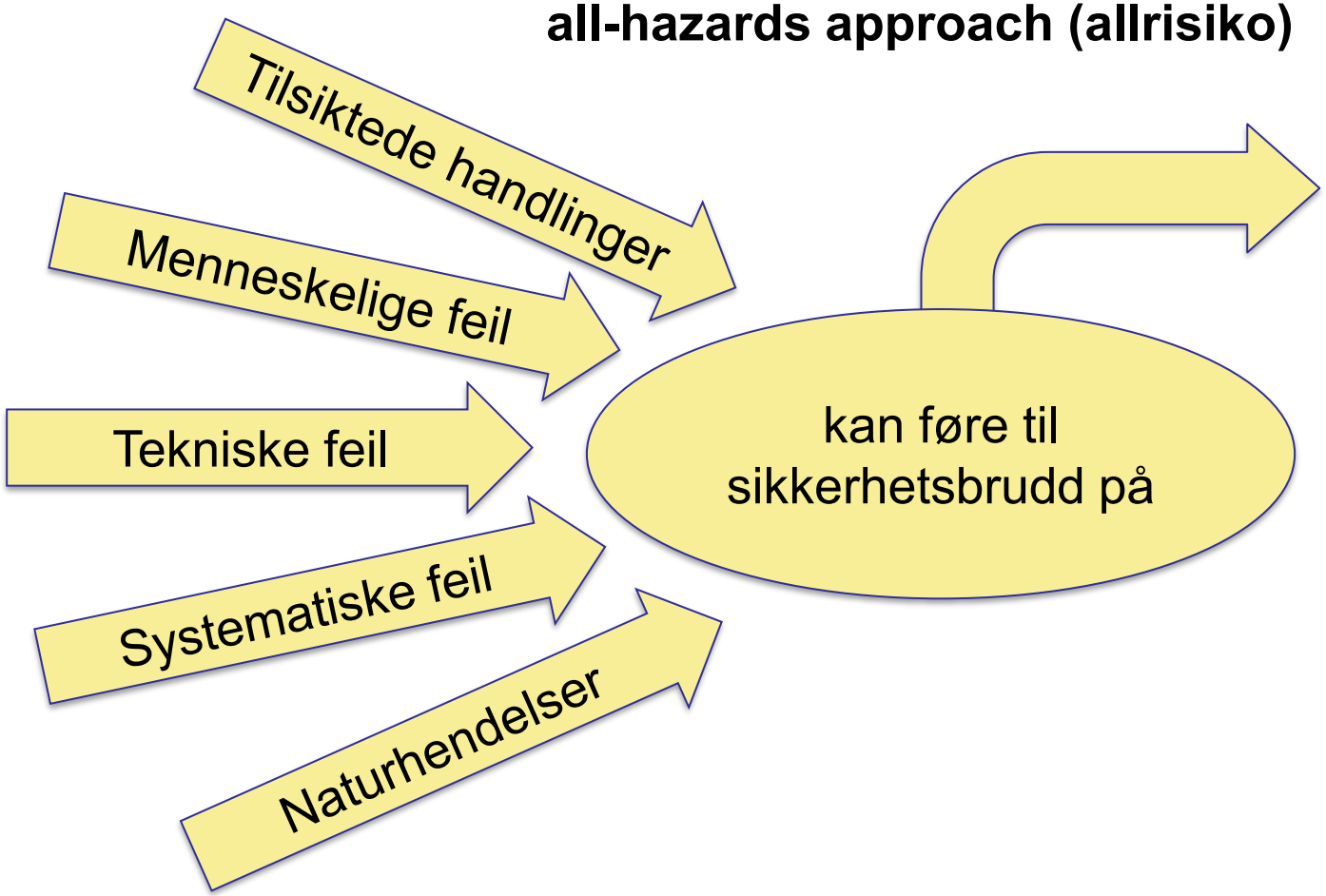
NIS2-direktivet (Network and Information Security nr. 2)



- NIS2-direktivet gjelder fra 2023, og erstatter det tidligere NIS-direktivet fra 2016. Da Norge og EØS-landene allerede har innlemmet det forrige NIS-direktivet som «lov om digital sikkerhet», betyr det at Norge også vil innlemme NIS2, som vil skje gjennom en ny eller endret lov. Elementer i NIS2 er blant annet:
  - skape samarbeid om en europeisk infrastruktur for håndtering av cyberkrise gjennom EU-CyCLONe (European Cyber Crises Liaison Organisation Network)
  - harmonisere sikkerhetskrav og rapporteringsplikter
  - sørge for at medlemslandenes cybersikkerhetsstrategier dekker sikkerhet i leveransekjeder, sårbarhetshåndtering, sikkerhet i kjernenettet og generell digital sikkerhetskultur og cyberhygiene
- Lov om digital sikkerhet (NIS2-direktivet) er en parallell til sikkerhetsloven, som skal gjelde for alle virksomheter med kritiske funksjoner, ikke bare «grunnleggende nasjonale funksjoner»

Fokus for EU direktiv NIS2 (Lov om digital sikkerhet) (Network and Information Security)

all-hazards approach (allrisiko)



KIKS: kritisk infrastruktur og kritiske samfunnsfunksjoner

- Nettverks- og informasjonssystemer (i KIKS), brukere av slike systemer og andre personer
- Personvern
- Liv og helse, miljø
- Grunnleggende nasjonale funksjoner (GNF)
- Nasjonale sikkerhetsinteresser

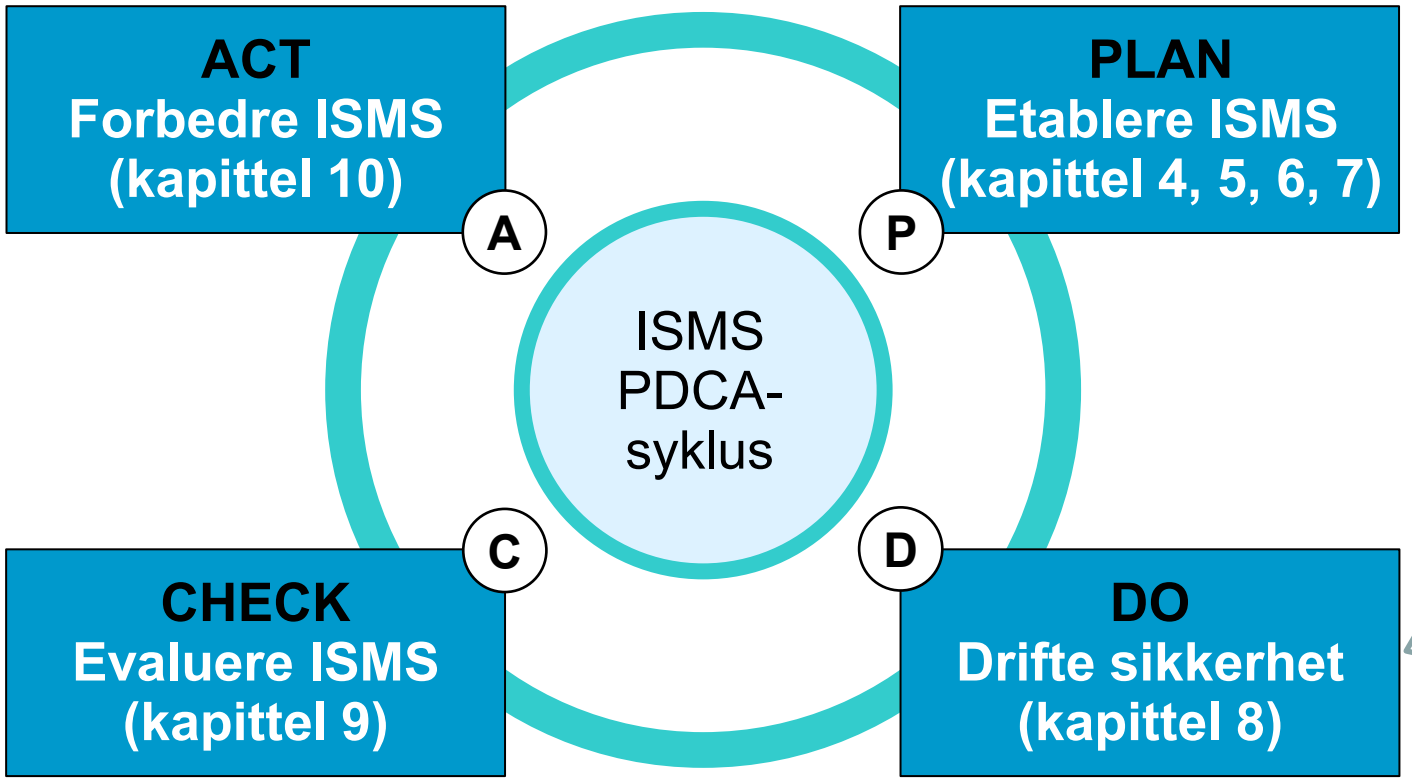
Lov om digital sikkerhet (digitalsikkerhetsloven)

- **Formål** (§ 1)
- Loven skal bidra til å sikre grunnleggende krav til digital sikkerhet i virksomheter med særlig betydning for samfunnet ved å forebygge, avdekke og motvirke uønskede hendelser i nettverks- og informasjonssystemer som brukes for å levere samfunnsviktige tjenester og digitale tjenester. Loven skal også legge til rette for sikkerhet i IKT-produkter, IKT-tjenester og IKT-prosesser.

ISO/IEC 27001 ISMS - krav

Krav til styrings-/ledelsessystem for informasjonssikkerhet

- ISO/IEC 27001 beskriver krav til ISMS, som kan struktureres og visualiseres med PDCA-syklusen.
- Konkret styrking av cybersikkerhet med valg og implementering av sikkerhetstiltak skjer i DO-fasen.
- God styring av DO-fasen krever oppfyllelse av kravene i de andre fasene.



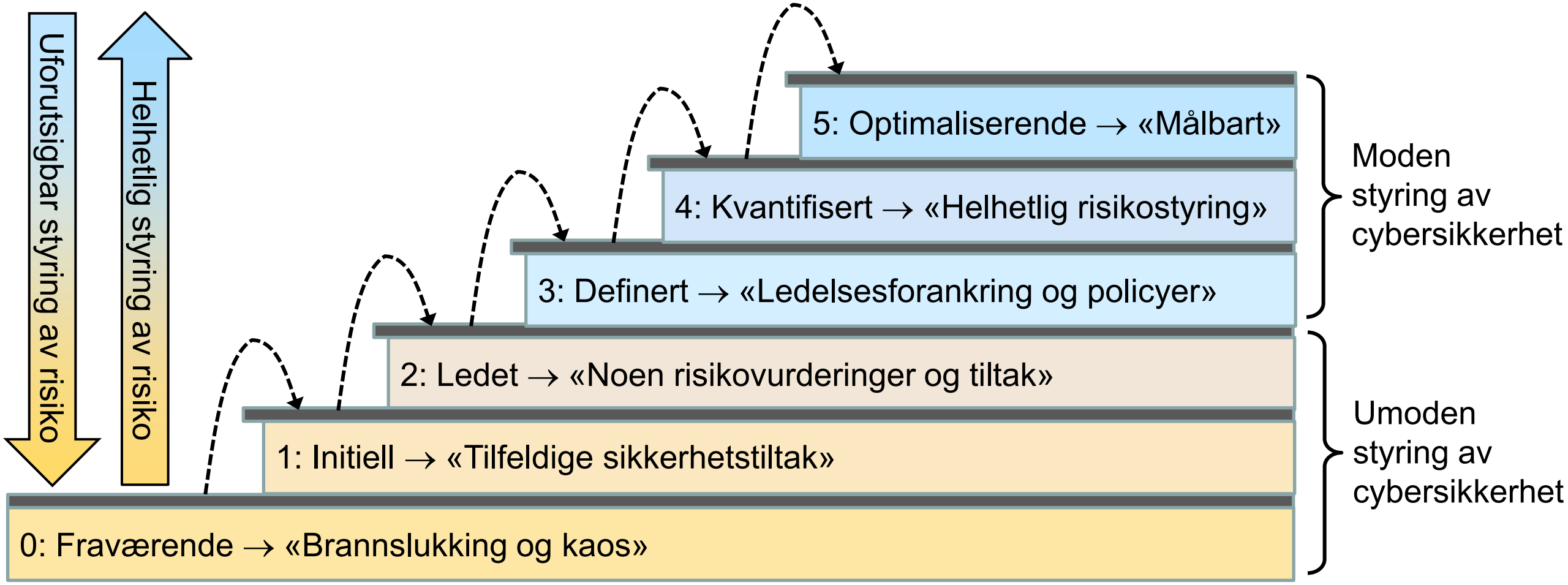
Implementere og drifte sikkerhetsfunksjonene fra NSMs grunnprinsipper for IKT-sikkerhet (alternativt fra NISTs CSF)

1. Identifisere og kartlegge
2. Beskytte og opprettholde
3. Oppdage
4. Håndtere og gjenopprette

ISMS/LSIS/Internkontroll - Kjært barn har mange navn.

- Styringssystem (ISMS) og ledelsessystem (LSIS) for informasjonssikkerhet har samme betydning.
- Digdir benytter betegnelsen «internkontroll» omtrent med betydning ISMS/LSIS.
  - Internkontroll/ styringssystem/ ledelsessystem for informasjonssikkerhet | Digdir
- Privat og statlig sektor (NSM) benytter vanligvis begrepet ISMS (Information Security Management System), oversatt som «styringssystem for informasjonssikkerhet».
- Lover og forskrifter benytter begrepet «styringssystem for informasjonssikkerhet» f.eks. eForvaltningsforskriften.
- Standard Norge oversetter ISMS som LSIS (Ledelsessystem for informasjonssikkerhet).
- Begrepet LSIS brukes ofte i undervisningssektoren i Norge, bl.a. på UiO.
- Se beskrivelse av UiO sitt LSIS:
  - https://www.uio.no/tjenester/it/sikkerhet/lsis/
- Begrepet LSIS benyttes av noen andre statlige virksomheter, bl.a. Sykehuspartner.
- Merk at internkontroll også kan bety «intern sikkerhetsrevisjon».

Modenhet i styring av cybersikkerhet
CMMI: Capability Maturity Model Integration



KI og Cyber



AI has crossed a threshold – what Claude Mythos means for the future of cybersecurity

Published: April 23, 2026 6.03pm CEST

The independently tested capabilities of Claude Mythos Preview have got the financial sector worried. Jinda Noipho/Alamy

The limit of what artificial intelligence can achieve, known as frontier AI, has crossed another threshold. AI can now plan and execute sophisticated cyber operations with minimal guidance at speeds far beyond human capability.

Author



Gerald Mako
Research Affiliate, Cambridge Central Asia Forum, University of Cambridge

KI og cyber



Offensiv KI

- Deepfakes
- Finne (nulldags)sårbarheter
- Produksjon av skadevare
- Angrepsautomatisering



Defensiv KI

- Inntrengningsdeteksjon
- Analyse av skadevare
- Digital trusseletterretning
- Hendelsesrespons



Sårbarheter og angrep mot KI

- Forgiftning av læringsdata
- Forurenset læring
- Leveransekjederisikoer
- Synsbedrag
- Lekkasje
- Injeksjonsangrep

TREND | Business

Unusual CEO Fraud via Deepfake Audio Steals US$243,000 From UK Company

September 05, 2019

An unusual case of CEO fraud used a deepfake audio, an artificial intelligence (AI)-generated audio, and was reported to have conned US$243,000 from a U.K.-based energy company. According to a report from the Wall Street Journal, in March, the fraudsters used a voice-generating AI software e[illegible]xecutive of the company's Germany facilitate an illegal fund transfer.



InSight Crime

Menu

News | Cybercrime

Fraud Groups Use Deepfakes to Enhance Imitation Scams in Peru

by *Gavin Voss*
21 Jul 2023



asia financial NEWSLETTER

AI

$600,000 'Deepfake' Fraud Heats Up AI Debate in China

May 22, 2023

"He chatted with me via video call, and I also confirmed his face and voice in the video. That's why we let our guard down," the victim said



Gadgets 360
An NDTV venture

HOME | GUIDE | AUTO | NEWS | REVIEWS | FEATURE
MOBILES | TELECOM | HOW TO | GAMING | ENTERTAINMENT

Home > Cryptocurrency > Cryptocurrency News > Crypto Scammers Using A

Crypto Scammers Using AI Deepfakes to Spoof KYC Verification on Exchanges, Binance Security Chief Says

Deepfakes are artificially generated photos or videos designed to can convincingly replicate the voice as well as facial features of an individual.

Written by Radhika Parashar, Edited by David Delima | Updated: 24 May 2023 15:13 IST

Share on Facebook | Tweet | Snapchat | Share | Reddit | Email | Comment | Google News



CNN Business Watch Listen Live TV Sign in

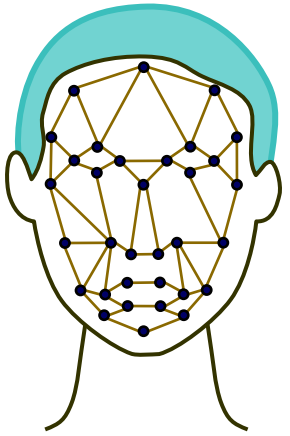
British engineering giant Arup revealed as $25 million deepfake scam victim

By Kathleen Magramo, CNN
3 minute read · Updated 4:53 AM EDT, Fri May 17, 2024





State-of-the-art Deepfake



- Deepfake av Mette Frederiksen statsminister i Danmark, laget av Morten Messerschmidt i Dansk folkeparti
  - https://twitter.com/MrMesserschmidt/status/1783882247323492725
- Deepfake av Fredrik Solvang
  - https://www.nrk.no/norge/ki-stunt-i-debatten-_-deepfake-av-fredrik-solvang-pa-direkten-1.16901230
- Olga Loiek stjålet avatar i Kina
  - https://www.youtube.com/watch?v=3FQSFnZpsqw
- Microsofts VASA-1 (Visual Affective Skills Avatar)
  - https://www.microsoft.com/en-us/research/project/vasa-1/

  «We have no plans to release an online demo, API, product, additional implementation details, or any related offerings of VASA until we are certain that the technology will be used responsibly and in accordance with proper regulations.»



Deteksjon av deepfake

- Deteksjon avhenger av at det ser **uekte** ut
- Deepfake lages for å se ekte ut
- Uløselig problem

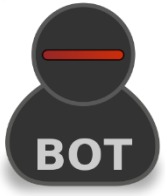


- Kryptografisk autentisering av video?



Produksjon av skadevare

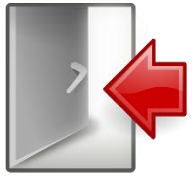


- Datavirus 
- Løsepengevirus 
- Spionvare 
- Bott-program 

- Exploit 

- Makro-virus 

- Trojaner 
- Dataorm 
- Rootkit 

- Bakdør 
- Skadelig JavaScript 

- Logisk bombe 

Angrepsautomatisering





Inntrengningsdeteksjon

Defensiv

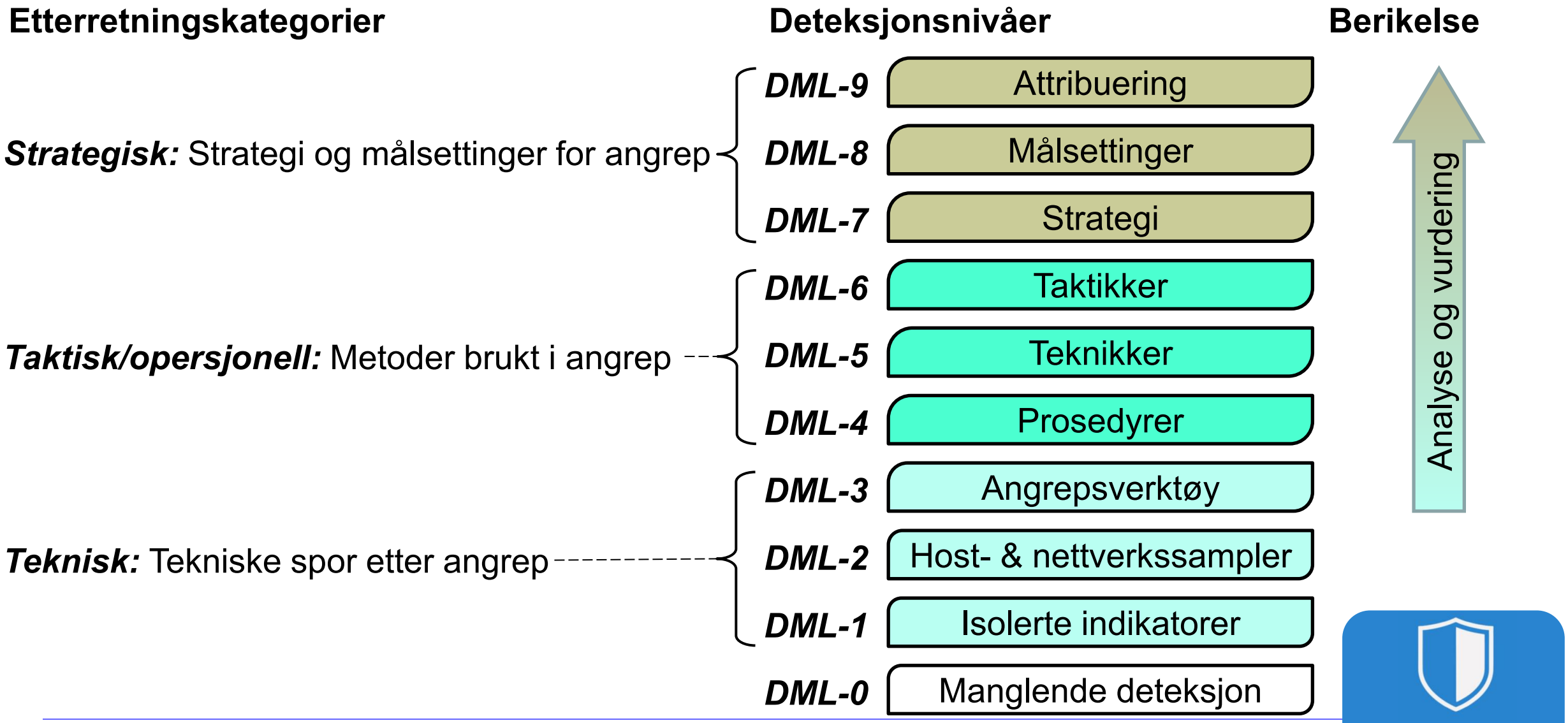


Analyse av skadevare

Defensiv



Digital trusseletterretning

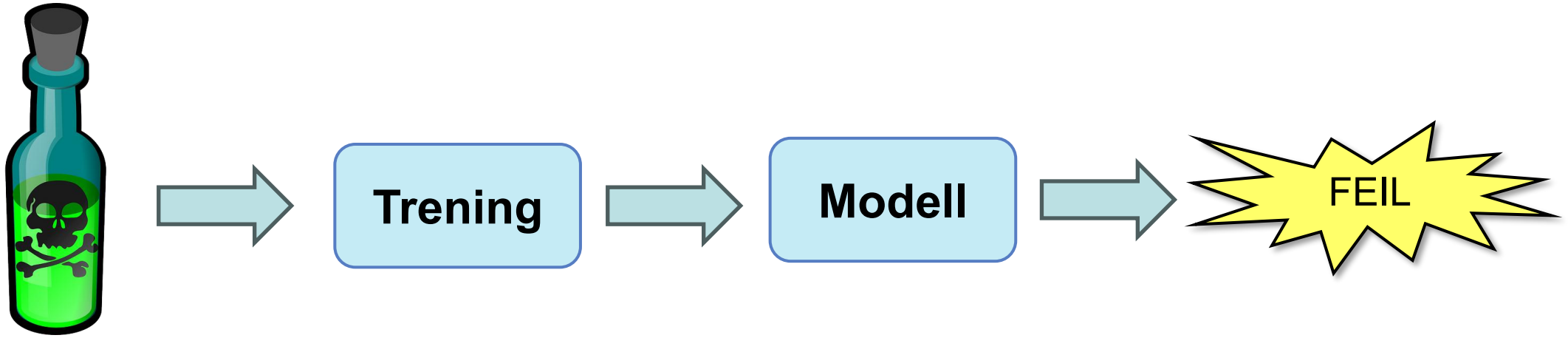


Hendelsesrespons

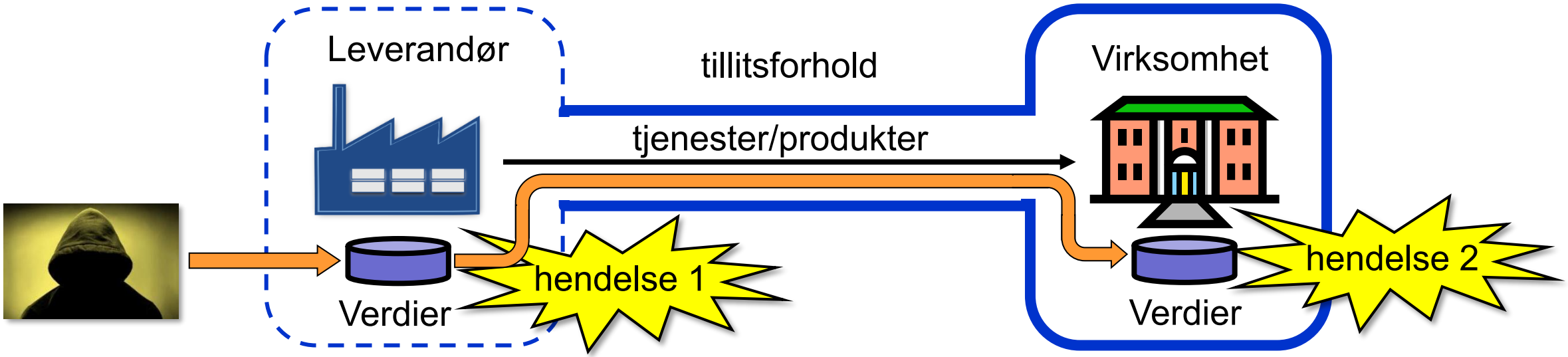
Defensiv



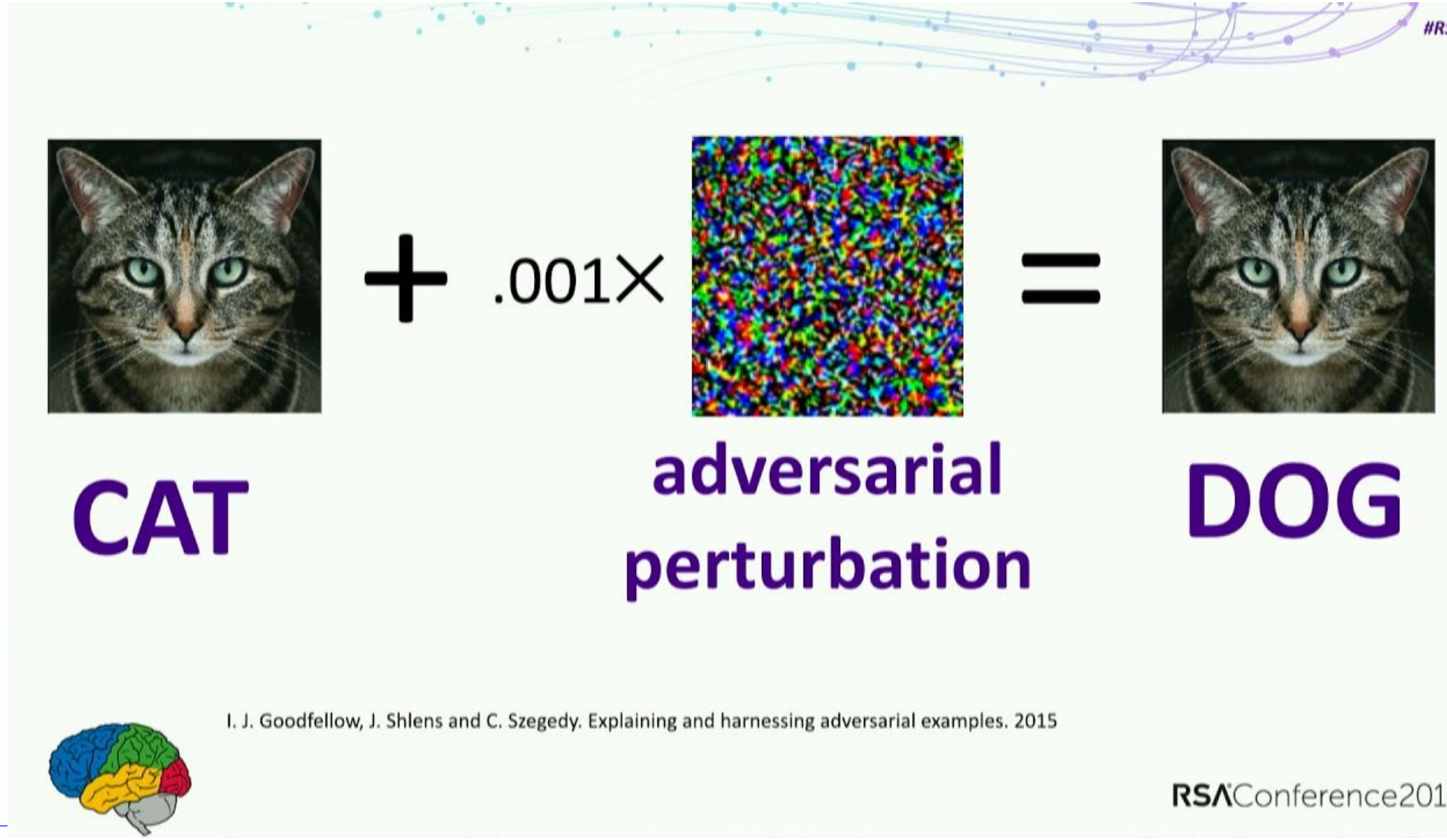
Forgiftning og forurensing av læringsdata



Angrep på leveransekjeden



Synsbedrag


#RSAC
.001×
CAT
adversarial perturbation
DOG
I. J. Goodfellow, J. Shlens and C. Szegedy. Explaining and harnessing adversarial examples. 2015
RSAConference2019

Synsbedrag



Human: 100.0 % stop sign
Machine: 99.7 % stop sign



Human: 100.0 % stop sign
Machine: 0.9 % stop sign

Jailbrake, lekkasje og injeksjonsangrep



Slutt på presentasjonen