



Putins neste trekk

HÅVARD GULLDAHL, journalist
INGHILD ERIKSEN, journalist
BENJAMIN FREDRIKSEN, journalist
FRIDA ANNETTE HELSETH STRØMME, designer
BJØRNAR DERVO, kartograf
JOHANNES ODLAND, utvikler

Data-SKUP 2026

Innsendere

Benjamin Fredriksen
Inghild Eriksen
Håvard Gulldahl
Frida Annette Helseth Strømme
Bjørnar Dervo
Johannes Odland

Ledelse

Redaktør Robin Mortensen
Reportasjeleder Gyda Katrine Hesla
Prosjektleder Marianne Orvik

Samarbeid

Danmarks Radio
Niels Fastrup
Lisbeth Quass

Sveriges Television
John Granlund
Simon Krona

Delfi Meedia
Kaur Maran
Holger Roonemaa

Barents Observer
Olesia Krivtsova
Georgii Chentemirov

Takk til

Erlend Koppergård, Ingrid Hjellbakk
Kvamstø, Beth Mørch Pettersen,
Johan Ante Utsi, Carl Gøran Larsson,
Bernt Olsen, Per Arne Staib, Hans
Ludvig Andreassen, Thomas
Sørensen, Knut Sætre, Caroline
Rugeldal, Rolf Jakobsen, Marina
Fridman, Thomas Nilsen

Marko Eklund

Publisert på [NRK.no](https://nrk.no) og NRKTV i juni og august 2026

NRK Undersøkende redaksjon nord

Innledning.....	3
Metodebruk i publikasjonene.....	4
Mottakelse og reaksjoner.....	5
AI-assistert koding – slik gjorde vi det etterprøvbart.....	5
En To Tre verktøy for å jobbe med satellittdata.....	6
AlphaEarth.....	7
Dynamic World.....	9
Lysanalyse.....	9
Hva kunne Yandex sine sladder si om russisk krigsøkonomi?.....	10
Svakheter.....	11
Analyse.....	12
Streisand-effekten.....	12
Design kan drive datastrategien.....	13
Lenker til publiserte saker.....	14

Innledning

I starten av 2025 hadde vi en rekke samtaler som etter hvert ble til dette prosjektet. Vi opplevde en dissonans mellom folk og forsvarskilder om hvor farlig situasjonen var for Europa.

Etter hvert som tiden dempet sjokket over Russlands invasjon i Ukraina i 2022, var det i starten av 2025 blandede signaler om Norges egen situasjon.

For eksempel skrev Etterretningstjenesten i sin åpne trusselvurdering på starten av året at vel hadde Russland «store planer for å utvide landmakten, herunder dens tilstedeværelse langs hele grensen til NATO», men tjenesten mente at myndighetene hadde liten evne til å realisere planene: «Moskva forsøker å gi inntrykk av styrke...», og videre at «det er sannsynlig at Russland tidligst kan realisere planene fullt ut fem til ti år etter at krigføringen i Ukraina opphører.»¹

Dansk etterretningstjeneste vurderte det annerledes. Da de på samme tid publiserte sin vurdering om krigsfaren skrev de at det var «risiko for krig i stor skala om få år», og satte en tidsskala på to til fem år². Det skapte bølger internasjonalt, men kritikk³ og avvisning⁴ her på berget.

Vi diskuterte dette med danske kollegaer i Danmarks Radio, som vi hadde jobbet med tidligere. Vi hørte fra kollegaer i Baltikum at det ble rullet ut spanske ryttere i bjørkeskogen, sperringer mot stridsvogner på grensa mot Russland. Hvem hadde mest rett – norsk eller dansk etterretning? Kunne norsk etterretning ha bommet på tempoet? Og hva så hvis de rustet opp. Var Norge noe mer truet av den grunn?

Journalister fra fire land begynte å jobbe med saken våren 2025. Vi samarbeidet med Danmarks Radio, Sveriges Televisjon og estiske Delfi Meedia. Målet var å undersøke nettopp den utmeldte opprustingen: var den reell, og hvor raskt kunne det gå?

Artiklene var de mest leste gravesakene på [NRK.no](https://www.nrk.no) per september 2026.

Prosjektet viser at samarbeid, nysgjerrighet og tydelige rammer for datahåndtering, datasett og programkode har gitt gode resultater.

Vi har erfart at noen ganger blir ting bedre av at man gjør feil først. Vi har opplevd at *visualisering* kan løse opp i fastlåste hypoteser og kodeproblem. Og oppi alt dette har vi utviklet en robust metode for KI-assistert koding som vi har med oss til alle nye prosjekter.

¹ https://www.etterretningstjenesten.no/publikasjoner/fokus/fokus2025_innhold/Fokus25-Kap1

²

https://www.fe-ddis.dk/globalassets/fe/dokumenter/2025/trusselsvurderinger/-20250209_opdateret_vurdering_af_truslen_fra_rusland_mod--.pdf

³ <https://www.vg.no/nyheter/i/LMX8ep/dansk-etterretning-risiko-for-storskalakrig-i-europa-i-loepet-av-faa-aar>

⁴ <https://www.dagbladet.no/nyheter/uansvarlig/82669484>

Metodebruk i publikasjonene

For å hjelpe oss med å forstå struktur, konvensjoner og planene for det russiske militæret, hyret vi inn en tidligere finsk etterretningsveteran. Han hadde tidligere vært utstasjonert i Moskva og hadde jobbet med militæranalyse for det finske forsvaret, nå frilanset han for nordiske medier.

Basert på sin inngående kjennskap til russisk militære, valgte han ut en rekke militære områder som vi skulle følge med på fra 2025 til 2026. Han mente de ville være eksempler på utviklingen i russisk opprusting. I hans metodekasse lå både fugl og fisk: Han skulle både følge med på kommersielle, høyoppløselige satellittbilder for å se om infrastruktur ble bygget eller utstyr ble ført til. Samtidig fulgte han med på åpne kilder fra bakkenivå som russiske nettsteder, nyhetsmeldinger og sosiale medier fra militæret og lokalsamfunnene. Akkurat som her hjemme blir det ofte omtalt lokalt i Russland hvis det skal bygges noe stort. Eksempler på dette kan være flere arbeidsplasser i byggebransjen, eller varsler om stans i vanntilførsel fordi en ny militærleir til noen tusen soldater er under bygging.

Hans konklusjoner la grunnlaget for den første av de tre sakene. Her dokumenterte vi at russisk militær opprusting var i full gang og i et høyere tempo enn hva vi har i Norge og NATO. Vår ekspert regnet ut at Russland bygde opp en landstyrke som kunne bli tre til fire ganger større enn tilsvarende på norsk og NATOs side. Et eksempel var fra Petsjenga, rundt 1 mil fra den norske grensa. Til denne saken la vi ned en betydelig innsats i å lage alternative datasett og utvikle nye verktøy. Formålet var å kontrollere og belyse ekspertens analyse.

For å jobbe med intervju av russiske kilder utviklet vi ulike metoder for automatisert oversettelse av tekst⁵ og lyd⁶ fra russiske kilder, uten å bruke skyløsninger. Slik kunne vi gjøre transkripsjon og analyse mens vi var på farten, og samtidig trygt for kildesensitivt materiale.

Den siste saken i trilogien baserte seg på et egenutviklet datasett over områder som var skjult og pikselert på russiske kart hos karttjenesten Yandex. Som første mediehus i verden viste vi et komplett datasett for grenseområdet og en kvalitativ undersøkelse av hver enkelt av de over hundre utviskede områdene. I et forsøk på å skjule, har hemmelighetene paradoksalt nok blitt mer synlige. Journalistikken ledet oss til en russisk rettslig ordre som var forsøkt hemmeligholdt, men åpnet døren inn til hvordan den russiske stat trekker stadig flere sivile selskaper og fabrikker inn i den store krigsøkonomien for å holde opprustingen i gang.

⁵ Her anbefaler vi den åpne modellen argos.translate: <https://www.argosopentech.com> Verktøyet vårt for å oversette har vi gjort tilgjengelig her: <https://gist.github.com/havardgulldahl/2a7ef3c440d2f0d934c4139259a21cfa>

⁶ Whisper-Faster

Mottakelse og reaksjoner

Da funnene ble presentert for forsvarssjef Eirik Kristoffersen, slo han offentlig alarm⁷. Overfor NRK erkjente han at han ikke var fornøyd med dagens beredskap, og medga en prekær mangel på ammunisjon og materiell. I møte med NRKs dokumentasjon innrømte Kristoffersen at Russland i ytterste fall ville kunne være klar til konflikt etter to-tre år – en tidslinje verken Forsvaret eller norske myndigheter hadde uttalt tidligere.

Prosjektet flyttet debatten fra om Russland vil bygge seg opp igjen, til hvor raskt de nordiske landene må omstille eget forsvar. Avsløringene utløste en debatt om omstillingstakten på norsk side av grensen – spesielt knyttet til om oppbyggingen av Finnmarksbrigaden skjer raskt nok⁸⁹.

Den russiske ambassaden avviste avsløringene som grunnløs vestlig propaganda og hevdet at Russlands militære aktivitet i egne grenseområder utelukkende er defensiv og i tråd med internasjonale avtaler. Sakene vakte oppsikt internasjonalt og ble flittig sitert av både etterretningsmiljøer og store europeiske mediehus som et sjeldent og konkret innblikk i Kremls langsiktige krigsforberedelser.

AI-assistert koding – slik gjorde vi det etterprøvbart

I løpet av prosjektperioden har verktøyene for bruk av AI til analyse og programmering utviklet seg i rekordfart. I starten jobbet vi med å kopiere kodesnutter inn og ut av verktøyene, mens utover 2026 begynte vi å bruke verktøy som kunne planlegge og behandle mye kode. Store språkmodeller og generativ AI gir åpenbart mange store fordeler for å heve kvaliteten og effektiviteten til både prosess og resultat. Samtidig er problemene med hallusinasjon og lav etterprøvbarhet nesten uhåndterlige for grundig og etterprøvbar journalistikk. Vår løsning var å bruke generativ AI til å lage *verktøy som ikke bruker generativ AI*. Slik kunne vi få god hjelp til å utforske, teste og skrive kode, men koden i seg selv skulle gi samme resultat hver gang vi kjørte den.

AI-hjelp til koding åpner også for et enda større, prinsipielt problem. Hvis du bruker AI til å gjennomføre oppgaver du ikke selv har kunnskap til å gjøre, er du per definisjon ikke i stand til å finne feilene når de dukker opp. Derfor er vi nøye med å lage testkode hele tiden, og å utvikle mange testcaser hvor vi på forhånd lager formelle krav til hva og hvordan verktøyene skal svare.

Vår superprompt¹⁰, som skal styre agentene underveis, inneholder blant annet

- Unngå store skript og lange funksjoner – gjør det mulig for oss å kontrollere koden.
- Skriv separate skript for hver datakilde – lettere å oppdage og forstå feil

⁷<https://www.nrk.no/tromsogfinnmark/forsvarssjef-eirik-kristoffersen-mener-forsvaret-ikke-bygges-fort-nok-sammenliknet-med-russland-1.17912233>

⁸<https://www.nrk.no/tromsogfinnmark/russland-ruster-opp-forsvaret-mot-nord-europa--stortinget-vil-ha-raskere-forsvar-sloft-i-norge-1.17917602>

⁹ <https://www.nrk.no/tromsogfinnmark/forsvarstopper-vil-ha-finnmarksbrigaden-pa-plass-allerede-i-2030-1.17921164>

¹⁰ F.eks. CLAUDE.md, GPT eller skill

- Skill mellom rådata og behandlet data – gir etterprøvbarhet
- Skriv like mye testkode som ekte kode

Andre regler vi styrer etter

- Legg inn PDF med artikler eller eksempler på hvordan en algoritme eller metode skal løses, som KI kan bruke som utgangspunkt
- Aldri slipp agentene løs i et filsystem som holder sensitive dokumenter. Hold det adskilt.
- Ikke driv med manuell datavask. Lag skript som gjør det, da har du samtidig dokumentasjon
- Alle analyser skal kunne gjøres fra skrætsj før publisering.

En To Tre verktøy for å jobbe med satellittdata

En av metodene for å følge med på utviklingen av militær infrastruktur var å sammenligne høyoppløselige satellittbilder over tid. Det er en kjent og effektiv metode. Her er et eksempel.



Satellite images in 2021 and 2025 of Russia's Sputnik military base near the border with Norway. The recent image shows expanded storage facilities for military equipment.

PLANET LABS PBC

I faksimilen fra Wall Street Journal¹¹ 28. april 2025, ser vi at to høyresolusjons satellittbilder ligger side om side. Slik ser leseren med det blotte øyet at tre nye bygninger kom til en gang mellom da bildet til venstre og bildet til høyre ble tatt. Det var akkurat slike bilder vi også var på jakt etter, men fra hele NATO-grensa. Spørsmålet vi stilte oss rett etter å ha startet prosjektet, var dette: *Kan vi lage en robot som finner alle områder hvor infrastruktur bygges?*

Det ville kunne løse noen problemer vi satt med.

- Var vi for avhengige av én datakilde?
- Kunne vi gå glipp av viktig utvikling dersom vi bare gjorde dette manuelt?
- Var bildedekningen god nok? Vinteren i nord har med seg både mørketid og mange skyer, så i verste fall kan det gå flere måneder uten et godt bilde i vinterhalvåret.

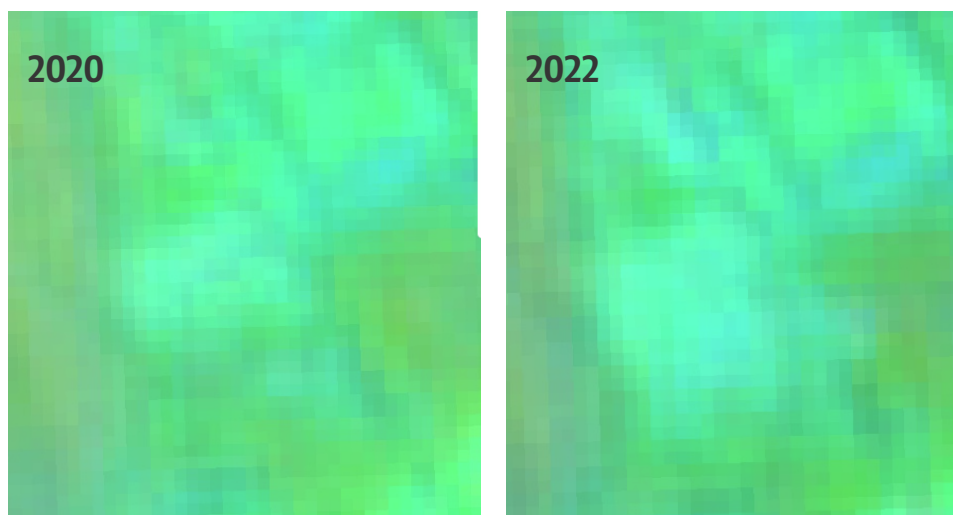
¹¹ <https://www.wsj.com/world/russia/russia-military-nato-europe-finland-ff53b912>

- Vi ønsket ikke å basere oss på enkeltbilder for å beskrive en trend. Det er metodisk svakt. Det er spesielt et problem hvis man også har dårlig bildedekning. Enn hvis de eneste bildene vi hadde var av de ekstreme tilfellene, og normalsituasjonen ikke var slik? Spesielt hvis bildene viste utstyrflytting var dette et viktig poeng.
- WSJ måtte angi et fem års tidsspenn, fordi det var de to bildene de hadde. Kunne vi bruke andre data for å forstå mer nøyaktig når endringer har skjedd, hvis vi ikke fant satellittbilder?
- Høyoppløselige bilder var dyre og krevde abonnement. Vi kunne dekke større områder dersom vi fant betalingsfrie datakilder.

Alle disse spørsmålene ledet oss ned en rute med mye koding, prøving og feiling. Alt i alt har vi jobbet med tre datasett, men bare ett av dem ble med i den ferdige journalistikken. Alle tre henter data gjennom Google Earth Engine¹², blant annet fordi da kan vi gjøre tunge kjøring på store datamengder, uten å måtte anskaffe dyre maskiner selv.

AlphaEarth

AlphaEarth var laget av Google Deepmind, som selv beskrev det som en «KI-modell som fungerer som en kunstig satellitt». Den var bygget opp fra mange ulike datakilder: optisk satellitt, radardata, laserkartlegging og mer. Første versjon var trent på over tre milliarder individuelle bilder. Modellen delte verden inn i 10x10 meter fliser, i 64 dimensjoner. AlphaEarth ga et årlig gjennomsnitt for billedata hver dimensjon i hver flis. Det gjorde den godt egnet til å se lange endringer, og dermed skulle sesongvariasjoner som snø og solmengde forsvinne. Det var relativt enkelt å komme i gang med å hente ut data¹³. Modellen inneholdt data som kunne beskrive endringene WSJ hadde vist bilder av – og at det hadde skjedd noe mellom 2020 og 2022.



Ser du forskjellen? Her er en visualisering av AlphaEarth sin data for samme utsnitt som WSJ har på bildene. Ser du nøye etter, kan du se de endringene AlphaEarth hadde plukket opp mellom årene.

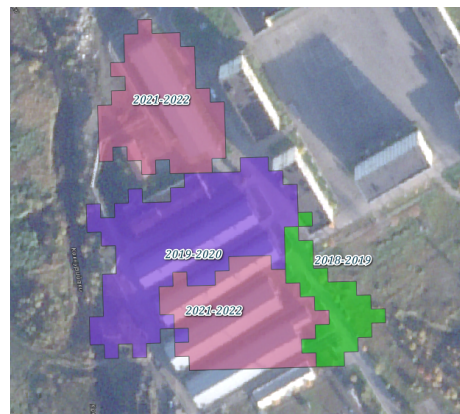
¹² Kom i gang med Google Earth Engine: <https://www.youtube.com/watch?v=I-wFYm4Hnhg>

¹³ Et forenklet skript kan du selv kjøre her:

<https://colab.research.google.com/drive/1exo7-MhPOVZxDOnNZt9cOFk5vvFTY4ji?usp=sharing>

Vår analyse med AlphaEarth ga oss en forklaring på hvordan infrastrukturen i militærbasen hadde utviklet seg år for år, selv uten at vi hadde visuelle satellittbilder tilgjengelig. Like viktig var at vi via Google Earth Engine matematisk kunne ta differansen mellom to års data for å beskrive utviklingen. Vi var ikke avhengig av visuell tolkning alene!

Ut ifra dette bygde vi et program som kunne ta et sett med lokasjoner (i form av en GPKG-fil¹⁴) og så gjorde en år-for-år sammenligning i modellen. Etter dette satt vi igjen med en ny GPKG-fil med mange lag, ett lag for hvert år i serien. Disse lasta vi inn i kartprogrammet QGIS for enkelt å sammenligne med billedata. For å kontrollere prosessen, testet vi programmet på en rekke steder hvor vi visste gjennom andre kilder at infrastrukturen hadde endret seg. Dersom verktøyet vårt var noe tross, så måtte det i alle fall finne disse. Slik kunne vi gå gjennom flere kjente steder (også WSJ-eksempelet) for å verifisere at vårt AlphaEarth-program kunne identifisere og isolere infrastruktur-endringer, år for år. For utvalgte områder hentet vi inn komplette tidsserier av satellittbilder fra Planet Labs for å verifisere at verktøyet vårt sine utregninger av endringer per år stemte.



Problemene meldte seg

På dette tidspunktet så alt veldig rosenrødt ut. Vi hadde brukt noen uker på å tilpasse parametrene i verktøyet slik at algoritmen klarte å analysere de kjente test-stedene vi hadde valgt ut. Til vår store glede fikk vi ofte ganske detaljerte svar, innenfor 10-20 meters oppløsning. Optimistisk meldte vi videre at vi snart kunne kjøre en analyse langs hele grensa. Men så begynte problemene å melde seg. Programmet var for ivrig.

- En rapportert *endring* betød ikke nødvendigvis at det var snakk om ny infrastruktur. Programmet plukket opp endringer på biloppstillingsplasser (flere eller færre biler), endringer i vegetasjon og et tilfelle av et tak som var malt i en ny farge. Programmet vårt kunne ikke skille mellom disse vidt forskjellige endringene.
- Dette kom av at hoveddatakilden i AlphaEarth utenfor de store byene var visuelle lavoppløselige satellittbilder, og ikke den typen signaler som kunne ha hjulpet med å skille ut infrastruktur fra annet (for eksempel lasermålinger og radar).

En manuell gjennomgang av alt programmet rapporterte ville vært enormt tidskrevende. Nå begynte vi å innse at vi kanskje var nødt til å lene oss på visuelle satellittbilder likevel. Hva var poenget med en fancy AlphaEarth-basert programvare dersom den bare rapporterte at *noe* hadde endret seg, men vi ikke visste om det var biler, maling eller myr?

¹⁴ En fil med geografisk data. Ligner på GeoJSON eller KML, men GPKG fungerer bedre når den er i et ledd av operasjoner.

Vi kontaktet en av forskerne hos Google som sto bak modellen og hadde publisert artikler om AlphaEarth. Hun bekreftet at modellen ikke «visste» hva som lå bak endringene. Det måtte man bruke andre midler til å finne ut av. Hun sa også at modellen var veldig ny, at det var mye de ikke visste selv om hvordan den fungerte. Derfor var det ikke så lett å svare på hvorfor programmet vårt ga ulike svar avhengig av hvilke år man sammenlignet med, annet enn å bekrefte at det var slik modellen fungerte.

Her begynte desperasjonen å melde seg. Problemet var at vi fikk rapportert mange endringer, men vi visste ikke hva de var. En idé meldte seg. Hva hvis man kunne bruke maskinlæring for å klassifisere endringene som AlphaEarth hadde funnet? Vi måtte ty til en gammel kjenning for trofaste lesere av Dataskup-posten.

Dynamic World

Vinneren av Dataskup 2023/2024 var prosjektet «Norge i rødt, hvitt og grått». Til grunn for prosjektet deres lå en maskinlæringsmodell som het «Dynamic World». Det var et ferdigtrent datasett for hele kloden, som kunne skille ulike typer vegetasjon samt bebyggelse fra hverandre. Gi Dynamic World et punkt, så svarte det om gress, myr, skog, vann eller bebyggelse var der. Vi utvidet programmet vårt til å kjøre klassifisering med Dynamic World. Vi testet den nye koden mot kjente tilfeller av oppgraderinger. Resultatene var nedslående. Selv om programmet nå ga et bedre resultat enn før, trengte det fortsatt mye manuell kontroll, og det var vanskelig å se at det kunne fungere i en frittstående prosess. Vi hadde måttet ta svært store forbehold i konklusjonene.

Vi hadde likevel kommet langt i forhold til målene vi hadde satt oss. AlphaEarth + Dynamic World lot oss oppdage utvikling selv når det ikke fantes satellittbilder, og ga en prosentvis antagelse av hva som hadde blitt bygget. Vi hadde klart å lage et verktøy som ga god kontroll på utvikling over tid – og kunne beskrive normalsituasjonen. Likevel gikk vi bort fra å bruke det videre. Det var to uoverkommelige problemer. Det første var at det var svært krevende å kontrollere funnene fra verktøyet. Høyoppløselige bilder måtte skaffes for å dobbelsjekke, og spesielt i de grisgrendte delene av Russland kunne det gå lang tid mellom gode bilder er tatt. I tillegg kunne det i mange tilfeller være vanskelig å sette skillet mellom militær og sivil oppgradering. Hvor mye bebyggelse rundt en militærleir er det riktig å regne med? Hva med utvikling på flyplasser som er både militære og sivile?

Vi kunne ha bestemt oss for at stikkprøver var godt nok, definert noen enkle regler eller prioritert å bruke mer tid på verifikasjon. Det gjensto som en mulighet for senere arbeid. Vi satt igjen med en forståelse av styrkene og svakhetene i AlphaEarth som kilde. Andre deler av det vi hadde lært skulle det vise seg at vi kunne gjenbruke.

Lysanalyse

Vi skjønte på dette tidspunktet at vi måtte finne en annen datakilde. Og nå hadde vi ikke tid til et komplisert rammeverk, nå måtte vi finne noe enkelt.

Løsningen var *lys*. NASA sine Black Marble-satellitter tar bilder av lys om natta. Det synes godt dokumentert at nattlys er en indikasjon på menneskelig aktivitet¹⁵. Kunne vi bruke denne informasjonen til å skape vårt eget bilde av hvor det var aktivitet?

Black Marble-datasettet viste seg å være akkurat enkelt nok til at vi kunne jobbe med det med den tiden vi hadde igjen av prosjektet¹⁶. For å bøte på eventuelle påvirkninger av nordlys, snø og is, valgte vi å bruke data fra sommermånedene. Vi beregnet også relativ heller enn absolutt endring i lys, for å være trygg på å kunne sammenligne på tvers av breddegrader. Vi brukte en logaritmisk skala på endring fordi betydningen av en viss mengde lys, dersom det fra før er mørkt, er større enn hvis det allerede er en del lys der. Vi hentet lysdata for 2022 og 2025 og beregnet differansen.

Endelig hadde vi fått til det vi hadde satt oss fore. Vi kunne vurdere konklusjonene fra ekspertene opp mot våre egne funn. Overordnet sett bekreftet lysanalysen ekspertens vurderinger og kildebruk. Det ga oss nye muligheter til hvordan vi presenterte utviklingen både i tekst og grafikk, ikke minst fordi vi kunne se utviklingen for hele områder av gangen. I den ferdige presentasjonen holdt vi oss likevel til de militære basene vi fulgte med på.

Hva kunne Yandex sine sladder si om russisk krigsøkonomi?

De første meldingene om at den russiske karttjenesten Yandex hadde begynt å viske ut militære områder på karttjenesten, kom på vinteren/våren 2025. En nyhetsmelding hos TASS¹⁷ i januar var første beskrivelse av at dette skjedde på russiske kart: Et raffineri måtte skjules på grunn av gjentatte ukrainske droneangrep. Senere hadde vi sett flere slike utviskinger på Yandex' kart. Vi la en ambisjon om å analysere Yandex' kart for å identifisere alle sladdene, researche dem grundig, slik at vi kunne snakke med eksperter om detaljene. Vi startet med noen enkle spørsmål.

- Hvor mange steder er fjernet?
- Hva ligger bak sladden?
- Hvis vi samler mange av dem - hva forteller helheten oss?

Rent praktisk begynte vi med å laste både Yandex og Google sine satellitt-kart som kartlag i kartprogrammet QGIS gjennom pluginen QuickMapServices¹⁸. Slik kunne vi enkelt inspisere områder som var fjernet fra Yandex – det var bare å flippe mellom lagene. Godt. Vi visste nå hvordan vi mennesker skulle kunne se og verifisere sladder. Men hvordan skulle vi lage et program for å skrape de samme områdene?

Vi prøvde ut en rekke forslag fra KI til hvordan vi skulle finne områdene med kode. Som vi nevnt i et tidligere kapittel, prøvde vi å hjelpe agentene med å vise til eksisterende kodebibliotek og publiserte

¹⁵ Lys og menneskelig aktivitet: <https://doi.org/10.1016/j.jag.2023.103627>

¹⁶ Her kan du selv se koden for å hente inn lysdata:

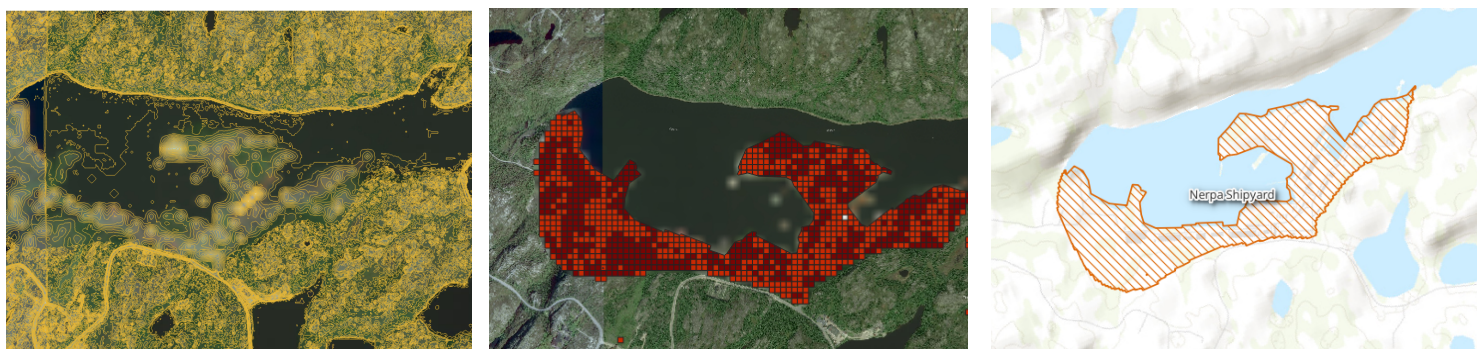
<https://colab.research.google.com/drive/1hjkyPclKWTlFhK-qNliaPBj9WhOc99L?usp=sharing>

¹⁷ https://web.archive.org/web/20260112101313/https://t.me/tass_agency/294471

¹⁸ https://plugins.qgis.org/plugins/quick_map_services/

algoritmer. Men i møte med Yandex-sladdene ble det alltid enten detektert for små, eller for store områder. Det hele virket uløselig. Så vi tok et skritt tilbake. QGIS er et avansert kartprogram, med all verdens algoritmer innebygget. Hvis vi kunne finne en metode for QGIS å isolere sladdene, var det i alle fall en start! Det viste seg etter hvert at et verktøy som heter «Contour» kunne skrus til for å gjøre jobben for oss. Det er egentlig laget for et helt annet formål, nemlig å tegne ut kurver fra en høydeprofil (tenk et bilde av et kart). Men tilfeldigvis fører sladdeprosessen til at de naturlige høydekonturene forsvinner fra satellittbildet, slik at dette verktøyet klarer å skille sladd fra ikke-sladd.

Etter å ha jobbet interaktivt med innstillingene i QGIS inntil vi fant parametre som fungerte på flere ulike sladder, lagde vi en plugin til QGIS som kunne analysere et bildeområde med *Contour* for å finne det sladdede området.



Fra rådata til objekt Ved å gjenbruke funksjonalitet i QGIS på en ny måte kunne vi lage en prosess for å finne alle sladdene på Yandex maps

Rent praktisk delte programmet bildeområdet opp i et rutenett og satte en verdi på glatthet per rute. Da var det mulig å sette en terskelverdi som gir færrest falske positive funn. Deretter lagde vi en prosesseringslinje som

1. Hentet inn bildedata fra Yandex fra et definert område av gangen
2. Delte bildedata opp i GeoTIFF (bilder med geografien innbakt som metadata)
3. Kjørte vår *Contour*-baserte plugin i parallell over geotiffene
4. Samlet sammen informasjonen inn i en GPKG-fil for inspeksjon av en journalist.

Svakheter

Programmet hadde noen svakheter. Det lette etter glatte, uniforme områder, og dette ga enkelte falske treff, for eksempel på bilder av skyer, snø, myr, tundra og store åkre. Derfor gikk journalistene gjennom alle figurene og undersøkte hvert enkelt treff. Enkelte av sladdene dekket så stort areal og så mange strukturer at det var vanskelig å lande på hva de var til for å skjule. Algoritmen måtte fortsatt renses av en journalist. Vi brukte QGIS også til dette, som med sine lag og allsidige geoverktøy kunne minne om arbeidsflyten i Photoshop. Til sist lastet vi opp område etter område til et prosjekt i Google Earth som vi ga alle journalister skrive-tilgang til.

Analyse

Vi hadde identifisert 120 sladdede områder langs NATO-grensa. Men hva dekket de? Journalister i fire land fordelte flekkene mellom seg, og gikk gjennom aviser, nettsteder, karttjenester og andre kilder for å avklare hva som lå bak sladden. Noen av kildene var: Wikimapia på russisk, OpenStreetMap sin «Se over elementer»-funksjon, kartlegging av bedrifter i den russiske forsvarsindustrien¹⁹, ukrainske droneangrep i Russland²⁰, russiske raffinerier²¹, sanksjonslister²² og relevant forskning.²³

Kategoriseringen viste et bredt omfang av sladdede områder: flyplasser, treningsleire, radaranlegg, fjellhaller for kjernevåpen, FSB-kontor, verft, forskningsinstitutter og fabrikker. I St. Petersburg var hele bykvartaler fjernet. Det var forvirrende at en hel del sladdede flyplasser og militærleire var nedlagt, forlatt og uten menneskelig aktivitet (vi gjorde en lysanalyse). Det var vanskelig å se det overordnede målet bak sladdene. Kilder vi snakket med, mente utvalget virket «nærmest tilfeldig», og at det faktisk kunne være et forsøk på forvirring.

Vi undersøkte hvordan tilsvarende prosesser forløp i andre land, deriblant den norske. Ifølge norsk lov skal skjermingsverdige objekter²⁴ skjermes. Interessant nok var det stor forskjell på hvor strengt de ulike kartaktørene, som Google, Apple og Bing, tolket dette, og graden av sladding varierte mye. Motvilje til sladding var også grunnen til at vi fant ut av hvorfor Yandex gjorde det.

Streisand-effekten

Den korte meldingen i januar 2025 om at Yandex hadde begynt å sladde, manglet både hvor og hvorfor. En ukrainsk notis satte oss på sporet av rettssaken bak. Via VPN hentet vi ut rettsavgjørelsen fra en russisk domstol i desember 2024. Yandex hadde nektet å sladde, så staten hadde gått til sak for å tvinge dem til å gjøre det. I dommen var det gjort en ufullstendig anonymiseringsjobb: En forkortelse for Rjazan oljeraffineri sto igjen i teksten. Dommen slo også fast at raffineriet hadde en militær rolle, og at Yandex derfor var nødt til å sladde: «AO RNPK arbeider i en uavbrutt modus for å sikre behovene til den russiske hæren og flåten...» Nå hadde vi en knagg å henge de uforklarlige sladdene på! Vi skjønnte at raffineriet hadde blitt satt på en gradert liste, etter den hemmelige «Ordre 138» fra industri- og handelsdepartementet – i praksis en mobiliseringsordre for russisk forsvarsindustri. Denne opplysningen ble kjernen i oppslaget, hvor vi vinklet på Ordre 138 og gjennom sladdene kunne vise at en rekke bedrifter og organisasjoner som leverte til forsvarsindustrien.

Etter publisering ble vi kontaktet av frivillige i OpenStreetMap. Etter avtale med dem la vi datasettet ut på nettet²⁵, slik at flere kunne jobbe med det. Avsløringen fikk god spredning internasjonalt. Det kunne se ut til at russiske myndigheter hadde utløst *Streisand-effekten*²⁶: I forsøket på å skjule

¹⁹ Dossier Center kartla over 6000 bedrifter knyttet til forsvarsindustrien

²⁰ Blant annet fra FIRMS

²¹ Caspian Policy

²² OpenSanctions

²³ https://www.thesimonsfoundation.ca/sites/default/files/MilitaryFootprintsInTheArctic_Final%2C%20March%202024.pdf

²⁴ NSM publiserte et kart over sensitive områder som ingen skulle vite om: <https://registrering.sensor.nsm.cloudgis.no/>

²⁵ Datasettet ligger på Github: <https://gist.github.com/havardgulldahl/3e92176bd5082f6961be012a36d62016>

²⁶ Barbara Streisand forsøkte i 2003 å forby spredning av et bilde av villaen sin. Da fikk det flere hundre tusen delinger

Rjazan-raffineriet og de 119 andre sladdene endte russiske myndigheter opp med å bekrefte den hemmelige ordren, koblingen til forsvarsindustrien og suksessen til de ukrainske angrepene.

Design kan drive datastrategien

Innad i NRK har prosjektet vært et samspill mellom datajournalistikk og presentasjon. Grafikere, kartograf og utviklere spilte på lag for å komme fram til det et resultat som kunne forstås av et bredt publikum. Samspillet var nytt for oss, og ga oss ny innsikt i strategier for å bygge riktige datagrunnlag. Et eksempel er at vi helt fra starten snakket om presentasjon av data.



Omriss og omhu Det var ikke nok å komme med et fancy datasett. Den visuelle prosessen hjalp med å unngå for kompleks presentasjon av lysdata. Det røde omrisset på sladden gjør det enklere å forstå abstrakte konsepter. Tredeling av endring i målt lys gjør det lettere å lese informasjonen.

Som tidligere nevnt, hentet vi opprinnelig inn lysdata for et veldig stort område av Russland. Det åpnet for muligheten til å vise «alt». Men i diskusjonen om hvordan dette skulle presenteres for publikum ble det klart at det var for mye. I en tidlig fase hadde vi definert en målgruppe (kvinne 30-40 år) og testet ulike eksempelsaker på et utvalg lesere. Tilbakemeldingene derfra var tydelig i retning av å forenkle datakildene og tallmaterialet som skulle være med. Derfor ble analysen av lysdata avgrenset til kun mindre områder. Et annet eksempel er at vi helt fra starten i arbeidet med å hente inn Yandex-sladder genererte skisser²⁷ til hvordan det kunne presenteres. Dette ga forståelsen av hvordan sladdene skulle skrapes for at det skulle fungere til slutt. Denne rapporten var også verdifull å ha i tilfelle Yandex Maps endret på sladdene, eller plutselig sluttet å virke.

I et overordnet perspektiv ser vi at samspillet mellom historiefortellere, datajournalister og designere har hjulpet hver av oss til å gjøre en bedre jobb. For selv om vi som undersøkende journalister har tilgang til mer avanserte verktøy enn noen gang, handler jobben ikke bare om *mer* data eller *mer* fancy analyse. Med stadig større og komplekse datamengder blir også behovet for tydeliggjøring og forenkling viktigere. Hvis dataene ser bra ut, vil flere forstå.

Datajournalistikk skal også nytes.

²⁷ Laget som interaktive HTML-filer, som ble zippet og sendt rundt. Slik kan mottakeren enkelt se det i nettleseren

Lenker til publiserte saker

1. <https://www.nrk.no/russland-ruster-opp-langs-grensa-mot-norge -nato-og-europa -overrasket-etterretningen-1.17889318>
2. <https://www.nrk.no/russiske-kilder-avslorer -dette-er-putins-plan-for-europa-1.17896258>
3. <https://www.nrk.no/ordre-138 -russlands-skjulte-forsvarsindustri-vokser-bak-sladder-pa-yan-dex--1.17930807>

Oppfølgere

- <https://www.nrk.no/tromsogfinnmark/forsvarssjef-eirik-kristoffersen-mener-forsvaret-ikke-bygges-fort-nok-sammenliknet-med-russland-1.17912233>
- <https://www.nrk.no/tromsogfinnmark/russland-ruster-opp-forsvaret-mot-nord-europa -stort-inget-vil-ha-raskere-forsvarsloft-i-norge-1.17917602>
- <https://www.nrk.no/tromsogfinnmark/forsvarstopper-vil-ha-finnmarksbrigaden-pa-plass-alle-rede-i-2030-1.17921164>
- <https://www.nrk.no/tromsogfinnmark/finnmarksbrigaden-foran-skjema -men-far-eldre-leopard-stridsvogner-1.17925103>
- <https://www.nrk.no/tromsogfinnmark/haersjefen-sier-alta-blir-viktig-for-at-oppbygging-av-finnmarksbrigaden-skal-lykkes-1.17926237>

NRKTV: <https://tv.nrk.no/program/KOID32001126>