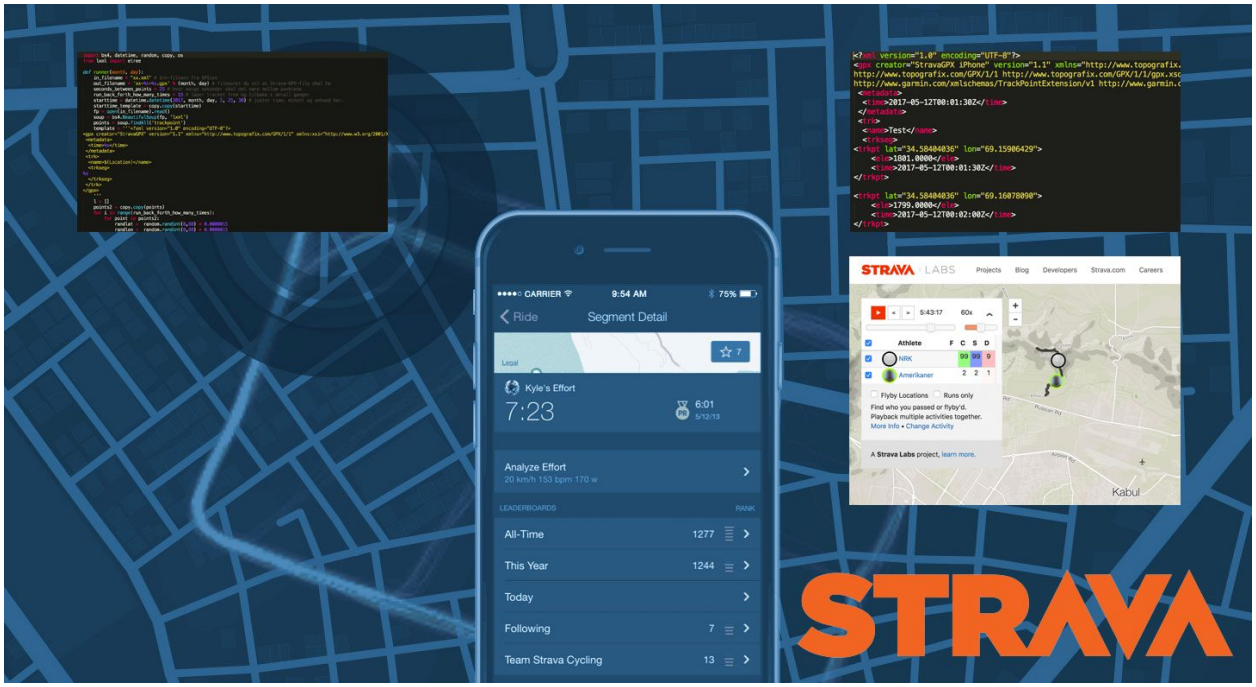
Metoderapport Data-SKUP 2018

Strava: Slik avslørte vi identiteten til militært personell i krigssoner



Prosjektnavn: Strava-avsløringen
Journalister: Henrik Lied (henrik.lied@nrk.no), Christine Svendsen (christine.svendsen@nrk.no) og Pedja Kalajdzic (pedja.kalajdzic@nrk.no)
Redaksjon: NRKbeta og NRK Nyheter
Dato: 30. januar 2018

1. Innledning

En virtuell joggetur lot oss avsløre identiteten til militært personell på sensitive utenlandsoppdrag.

Studenten Nathan Ruser oppdaget i januar 2018 at karttjenesten til Strava, en av verdens mest populære treningsapper, avslørte hvor det finnes aktive militærbaser rundt om på kloden – også på steder hvor militære makters tilstedeværelse ikke snakkes så høyt om.

Stravas karttjeneste samlet treningsrutene til alle deres millioner av brukere i et stort «heatmap», som gjorde det mulig å se omrissene til militærbaser og andre sensitive plasseringer verden rundt.

En rekke nasjonale og internasjonale medier omtalte problematikken, og også vi i NRKbeta var i ferd med å skrive en rask sak.

Men underveis fant vi ut at sikkerhetshullet i Strava skulle vise seg å være langt mer alvorlig:

Ved å programmatisk generere tusenvis av falske joggeturer kunne vi avsløre identiteten til en rekke soldater på oppdrag i utlandet, deriblant en nordmann.

2. Hypotese og prototype

2.1 Hypotese

Takket være en kollega som er relativt opptatt av både teknologi og trening, ble vi introdusert for en funksjon i Strava som heter Flyby.

Enkelt forklart er Flyby en tjeneste som lar deg se hvilke andre Strava-brukere som har jogget, syklet eller løpt i samme område som deg selv, på samme tid, i et interaktivt kart.

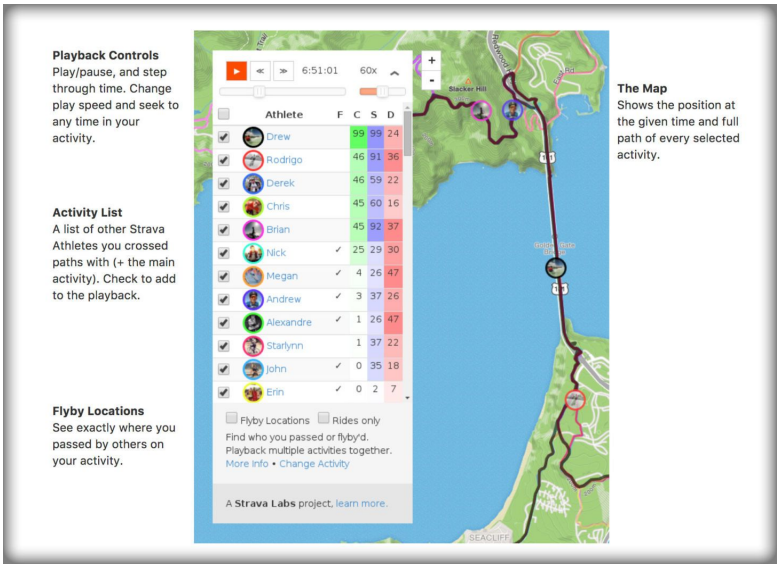
Dette dannet basisen for vår hypotese: Kan vi, ved å laste opp falske joggeturer til vår Strava-profil, finne vestlig militært personell i krigs- og konfliktsoner?

2.2 Research

Selv om mesteparten av Strava-bruken går via apper til iOS og Android, er det fortsatt mennesker som bruker gamle, "dumme" enheter som lager GPS-spor som filer man må hente ut fra enheten.

Siden Strava har et ønske om å kunne tilby sine tjenester til så mange mennesker som mulig, er disse litt gammeldagse dingsene støttet. Man kan laste opp såkalte GPX-filer, som inneholder GPS-informasjon om en tur, via Stravas web-grensesnitt.

Denne funksjonaliteten skulle vise seg å være instrumentell for å sjekke hypotesen vår.



***Figur 1**: Slik markedsfører Strava sin Flyby-funksjon.*

Flyby-funksjonen gir deg kun treff på andre brukere som har jogget *innenfor samme tidsrom*. Dette førte til at vi måtte tenke kreativt på hvordan vi skulle gjennomføre testen.

2.3 Prototype

Vi begynte med å analysere hvordan GPX-filene til Strava var strukturert, for å danne oss et bilde av hvor omfattende arbeidet vårt ville bli.

```
<?xml version="1.0" encoding="UTF-8"?>
<gpx creator="StravaGPX iPhone" version="1.1" xmlns="http://www.topografix.com
http://www.topografix.com/GPX/1/1 http://www.topografix.com/GPX/1/1/gpx.xsd ht
http://www.garmin.com/xmlschemas/TrackPointExtension/v1 http://www.garmin.com/
 <metadata>
  <time>2017-05-12T00:01:30Z</time>
 </metadata>
 <trk>
  <name>Test</name>
  <trkseg>
<trkpt lat="34.58404036" lon="69.15906429">
    <ele>1801.0000</ele>
    <time>2017-05-12T00:01:30Z</time>
</trkpt>

<trkpt lat="34.58404036" lon="69.16078090">
    <ele>1799.0000</ele>
    <time>2017-05-12T00:02:00Z</time>
</trkpt>
```

***Figur 2**: Et utdrag av en eksempel-GPX fra Strava.*

GPX-formatet er basert på XML, et dataformat som er ganske enkelt å lese. Noen minutter senere hadde vi god oversikt over hva vi måtte gjøre for å gjennomføre denne testen.

Prototypen ble et Python-program som genererte én fil per dag i året. Filene inneholdt 12 timer lange joggeturer for områder hvor vi antok at det ville være militær tilstedeværelse.

Vi satte i gang med utviklingen av et Python-script som:

- Tok en GPX-fil med en konstruert rute som input-fil
- Masserte strukturen i denne filen til å stemme overens med et format støttet av Stravas opplastingstjeneste
- Modifiserte strukturen slik at våre «fiktive løpeturer» ble ordentlig lange
- Laget 365 varianter av denne ruten – *en for hver dag* – som kan lastes opp til Strava
- For å dekke over et så stort område som mulig, brukte vi en *random-funksjon* for å skape litt variasjon i joggeturene våre. Hvert geografiske punkt fikk altså en liten forskyvning i posisjon.

```
import bs4, datetime, random, copy, os
from lxml import etree

def runner(month, day):
    in_filename = "xx.xml" # Inn-filnavn fra GPSies
    out_filename = 'xx-%s-%s.gpx' % (month, day) # filnavnet du vil at Strava-GPX-fila skal ha
    seconds_between_points = 25 # hvor mange sekunder skal det være mellom punktene
    run_back_forth_how_many_times = 15 # løper tracket frem og tilbake x antall ganger
    starttime = datetime.datetime(2017, month, day, 2, 25, 30) # juster time, minutt og sekund her.
    starttime_template = copy.copy(starttime)
    fp = open(in_filename).read()
    soup = bs4.BeautifulSoup(fp, 'lxml')
    points = soup.findAll('trackpoint')
    template = '''<?xml version="1.0" encoding="UTF-8"?>
<gpx creator="StravaGPX" version="1.1" xmlns="http://www.topografix.com/GPX/1/1" xmlns:xsi="http://www.w3.org/2001/X
 <metadata>
  <time>%s</time>
 </metadata>
 <trk>
  <name>$(Location)</name>
  <trkseg>
%s
  </trkseg>
 </trk>
</gpx>
    '''
    l = []
    points2 = copy.copy(points)
    for i in range(run_back_forth_how_many_times):
        for point in points2:
            randlat =  random.randint(0,80) * 0.0000015
            randlon =  random.randint(0,80) * 0.0000015
```

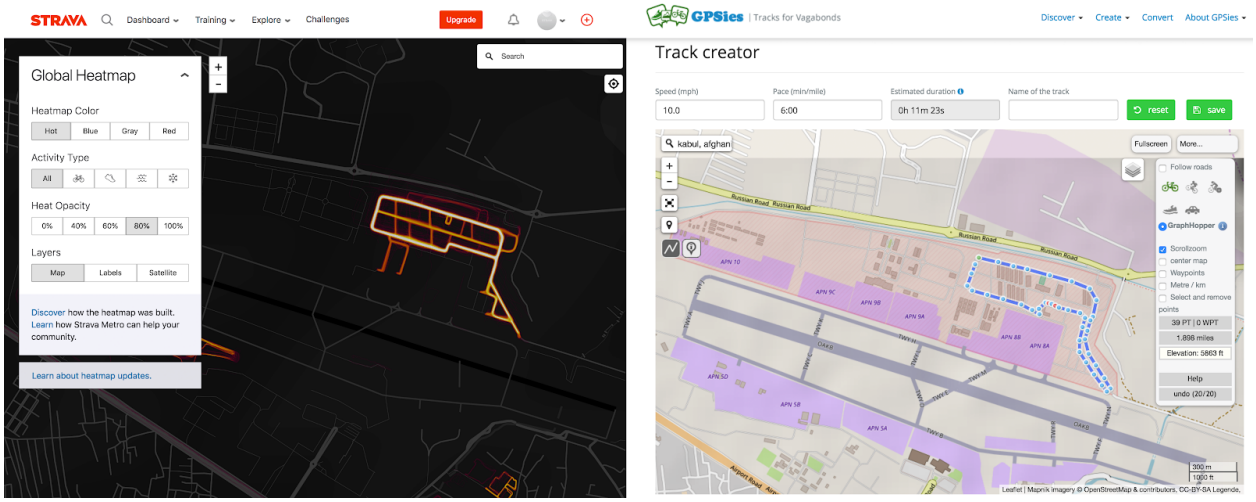
***Figur 3**: Et utdrag av Python-scriptet som genererte de falske løpeturene.*

Programmet er totalt rundt 60 linjer, og er skrevet i Python med litt hjelp av pakken BeautifulSoup for bearbeiding av XML-strukturen. Prototypen av scriptet ble utviklet på omtrent en halvtime, med små justeringer og forbedringer underveis.

Vi valgte oss ut Kabul som en første test, siden Afghanistans befolkning har lav mobildata-bruk. Dette øker sannsynligheten for at eventuelle spor som noen har lagt bak seg i Kabul, ikke kommer fra lokalbefolkningen. Samtidig vet vi at norske styrker i en årrekke har vært stasjonert i Kabul, så vi hadde forhåpninger om å finne nordmenn gjennom denne metoden.

Vi åpnet Stravas Heatmap (https://www.strava.com/heatmap) i en nettleser, og zoomet oss inn til Kabul. Da ser man umiddelbart en rekke områder som lyser opp.

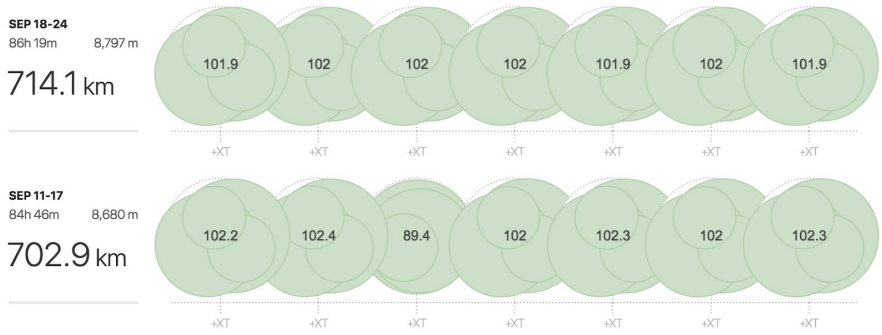
I et annet nettleservindu åpnet vi GPSies, en interaktiv nettside som lar deg tegne streker på et kart, og eksportere de som GPS-spor.



***Figur 4**: Ved å ha Stravas Heatmap åpent i ett nettleservindu, og GPSies i et annet, kunne vi lage et falskt GPS-spor som korresponderte med løpeturene til soldatene.*

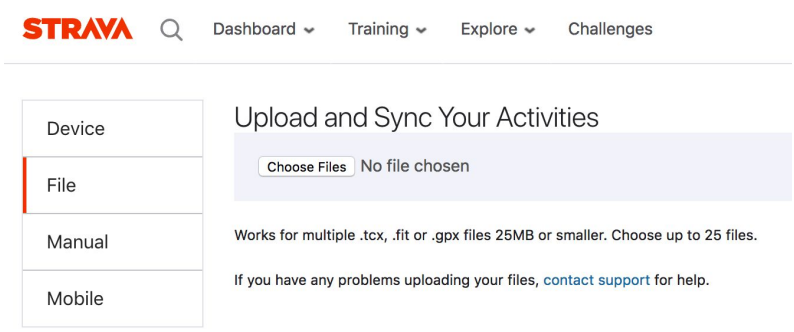
Vi prøvde så godt vi kunne å markere de samme områdene i GPSies, som de vi så i Stravas Heatmap. Da vi var fornøyde med sporet, eksporterte vi det som en Garmin GPS-fil, og kjørte filen gjennom Python-programmet vårt.

Da fikk vi 365 utgaver av det samme sporet, én for hver dag i 2017. Hvert spor ble strukket i tid, slik at det for Strava så ut som at vedkommende hadde gått kjempesakte gjennom Kabul. Turene ble cirka 12 timer lange. Gåturene ble lagt til ulike tider på døgnet, for å dekke opp for eventuelle problemer med tidssonekonvertering og utetemperatur. Vår hypotese var at det er mer sannsynlig at man treffer soldater på joggetur enten tidlig på morgenen eller sent på kvelden, siden temperaturene i Kabul fort kan komme opp i 40 grader.



***Figur 5**: En oversikt som viser antall kilometer vår falske Strava-bruker har løpt i september 2017.*

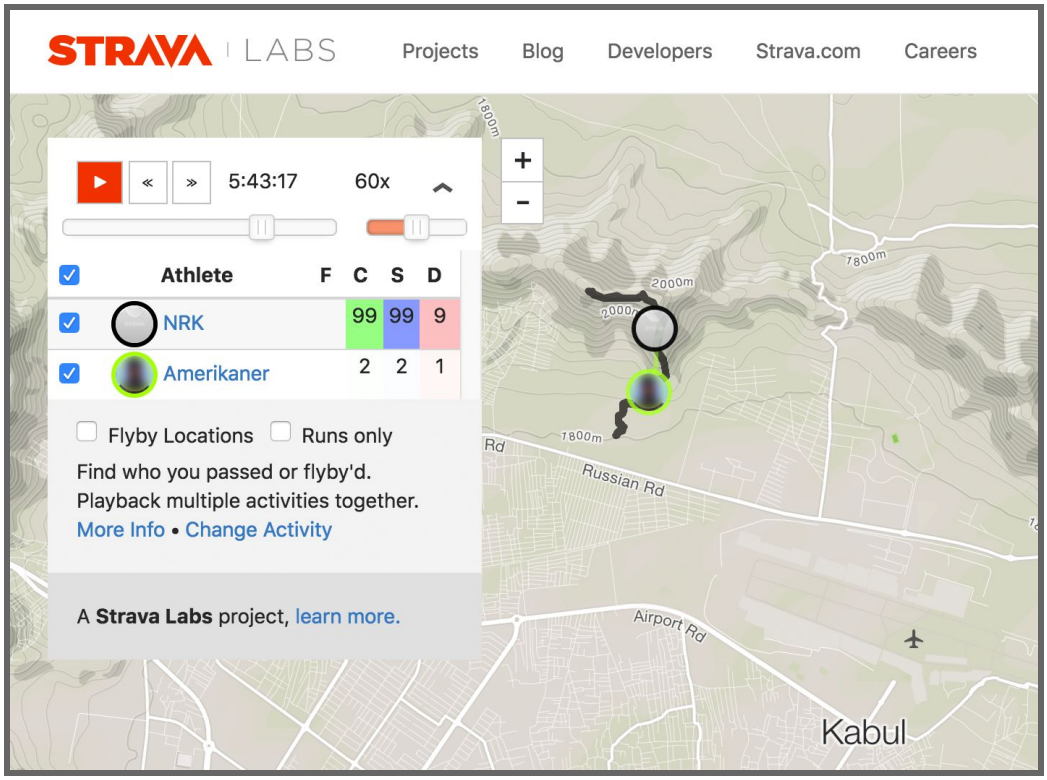
De 365 filene lastet vi så opp til en fiktiv Strava-profil. På grunn av begrensninger i brukergrensesnittet, og problemer med den programmatiske opplastingstjenesten (API-et), fikk vi bare lastet opp 25 filer om gangen. Det ble med andre ord en del manuelt arbeid for å få filene inn i Strava.



***Figur 6**: Gjennom dette grensesnittet lastet vi opp batcher på 25 filer til vår Strava-profil.*

Da dette var unnagjort, gjenstod den repetitive delen av jobben: Vi måtte åpne hvert eneste av de 365 sporene i Strava Flyby, og manuelt sjekke hvorvidt det var noen andre Strava-brukere på joggetur i den aktuelle perioden.

Vår ferd gjennom Kabul viste at systemet fungerte. Vi fant flere amerikanske soldater som jogget rundt på militærbaser og andre eiendommer i Kabul by.



***Figur 7**: Her "traff" vi en amerikansk soldat på joggetur i Kabul. Vi har endret navnet til vedkommende til "Amerikaner" av hensyn til hans sikkerhet. Hans profilbilde er også blitt tilslørt.*

3. Veien mot publisering

Da vi fikk bekreftet at prototypen fungerte, ringte jeg, Henrik Lied, til Christine Svendsen i NRK Nyheter, og spurte om vi skulle samarbeide på saken. Svendsen har i en årrekke jobbet med utenriksstoff, og har god kunnskap om militære enheter i både inn- og utland.

For vår del var det mest interessant å finne norske soldater på oppdrag i utlandet, så vi kartla hvilke områder som hadde norsk militær tilstedeværelse.

Etter at vi hadde utarbeidet en liste, begynte vi det nitide arbeidet med å lage spor i GPSies som samsvarte med de varme linjene i områdene fra Stravas Heatmap.

Vi sjekket seks ulike lokasjoner, og lastet opp filer for både hele 2016 og 2017 på de aktuelle lokasjonene. Totalt lastet vi opp nærmere 5000 filer til Strava, og sjekket hver eneste fil for aktivitet.

Underveis masserte vi Python-scriptet til å generere en liste over alle Strava Flyby-turene som vi hadde lastet opp, slik at vi slapp å gå innom Stravas webgrensesnitt for hver gang vi skulle sjekke om turen gav noen treff.

På under 6 timer hadde vi utviklet hypotesen, laget programmet og sjekket en rekke militære installasjoner i flere ulike land.

I tillegg til en norsk soldat, fant vi i løpet av noen timer identiteten til personer fra:

- Danmark
- USA
- Frankrike
- Nederland
- Italia
- Storbritannia

Mange av soldatene var på kjente, offisielle oppdrag, mens andre var en del av militære spesialstyrker, som gjerne prøver å holde sin tilstedeværelse skjult.

Jobben med å kartlegge menneskene vi fant via Strava Flyby ble stort sett gjort med enkle Google- og SoMe-søk. Når vi fant en person som hadde jogget i samme område som oss, gikk vi inn på vedkommendes profil, og noterte oss alle detaljer vi kunne finne. Ofte var det bare navn og et profilbilde, andre ganger var det flere ledetråder, som hvilke grupper og lag vedkommende hadde tilhørighet til. Det var også mye verdifull informasjon i vennelistene, siden vi da kunne grave videre på vennenes profiler for å finne bosted, arbeidsplass og andre identifiserende detaljer.

4. Derfor var saken viktig

Militært personell er attraktive terror- og etterretningsmål, og utilsiktet lekkasje av sensitive data kan få dramatiske konsekvenser. At soldater på utenlandsopphold avslører hvor og når de er på joggetur, samt posisjoner og tidspunkt for når de beveger seg utenfor sikrede baser, er svært alvorlig. I flere tilfeller oppdaget vi at soldatene hadde glemt å skru av Strava etter endt treningsøkt, noe som gav oss god innsikt i hvor de patruljerte. Dette er informasjon både terrorister og fremmed etterretning er svært interessert i.

I Irak fant vi også sivile, vestlige mennesker som brukte Strava på både sykkel- og joggeturer. Disse er særlig eksponert for fare, da de ikke har det samme sikkerhetsopplegget rundt seg som militært personell.

Denne saken er et godt eksempel på hvorfor *digital sikkerhet* er et vanskelig felt. Til og med spesialsoldater på utenlandsoppdrag, som er topptrent i fysisk sikring og etterretning, tenker ikke over sporene man legger bak seg ved å bruke mobilapper. Mange tenker på Strava som en ren treningsapp, men faktum er at all aktivitet man bedriver i appen, synkroniseres til serverne hos en amerikansk, kommersiell tjeneste.

Det er flere måter soldatene kunne unngått å gå i baret på, blant annet ved å justere på personvernsinnstillingene i appen. Hvis de berørte soldatene hadde satt sikkerhetsinnstillingene til høyeste nivå, ville joggeturene deres ikke blitt en del av det offentlige Strava Heatmap.

Standardinnstillingene til Strava er at brukerne deler all informasjon, siden dette gagner selskapet virksomhet. Selv om Strava i all hovedsak brukes som et loggeverktøy for trening, må man ha i bakhodet at Strava er et sosialt nettverk. Et viktig aspekt ved digital sikkerhet er å sette seg inn i hva de ulike tjenestene man bruker lagrer og viser av data, og denne saken er en god illustrasjon på at personvernsinstillinger ikke er aktivt justert av brukerne.

I kjølvannet av Strava-saken, lanserte både amerikanske Department of Defense og det norske Forsvaret reviderte retningslinjer for bruk av tjenester og apper som registrerer personlige opplysninger og posisjonsdata.

Da vi publiserte saken, innrømmet Forsvarets Operative Hovedkvarter (FOH) overfor NRK at dette er problematisk, og at det er noe de vil ha mer fokus på fremover. Uken etter publisering kontaktet Forsvaret alle norske styrkesjefer i utlandet, med beskjed om at de skulle bevisstgjøre norsk personell på at de må følge retningslinjene for bruk av applikasjoner med stedstjenester.

5. Til slutt

Strava-avsløringen er ikke et langt og strukturert prosjekt, men heller et godt eksempel på at man kan oppnå stor gevinst ved å jobbe raskt og tverrfaglig. Uten kombinasjonen av god teknisk innsikt og grunnleggende programmeringsferdigheter, samt kunnskap om militære operasjoner, hadde denne saken aldri blitt noe av.

6. Oversikt over saker

Slik røper soldater fra Norge, Danmark og USA hvem de er og hvor de trener i krigssoner
https://www.nrk.no/urix/slik-roper-soldater-fra-norge_-danmark-og-usa-hvem-de-er-og-hvor-de-trener-i-krigssoner-1.13891513

How soldiers from Norway, Denmark and USA disclose who they are and where they exercise in war zones
https://www.nrk.no/urix/how-soldiers-from-norway_-denmark-and-usa-disclose-who-they-are-and-where-they-exercise-in-war-zones-1.13892695

Strava: Slik avslørte vi identiteten til militært personell i krigssoner
https://nrkbeta.no/2018/01/30/strava-slik-avslorte-vi-identiteten-til-militaert-personell-i-krigssoner

How we found the identity of military personnel using Strava
https://nrkbeta.no/2018/01/31/how-we-found-the-identity-of-military-personnel-using-strava/

Strava-avsløringen: Forsvaret kontakter alle norske styrkesjefer i utlandet
https://www.nrk.no/norge/strava-avsloringen_-forsvaret-kontakter-alle-norske-styrkesjefer-i-utlandet-1.13895557

7. Kode

Koden for å generere fiktive Strava-filer ligger på Github:
https://gist.github.com/nrkbeta/15345a9dc78578fc2174b939661ad0d0