EINDVERSLAG DIGITALE ZOEKTOCHT NAAR HET SPECIAL REPORT (2010)
Oktober 2022 – Juli 2023

Datum: 28 juli 2023

Inleiding
In juli 2021 heeft [redacted] (P-CDS) de opdracht verstrekt om het missiearchief Afghanistan 2005-2010 te verbeteren en te ontsluiten.[1] Initieel werd deze opdracht uitgevoerd door een zogeheten Afghanistan Archiefteam (A-team) bestaande uit vier oud-militairen die gemiddeld 12 uren per week voor dit project werden vrijgemaakt. In juli 2022 werd het projectteam verder uitgebreid met een drietal *data scientists* van Accenture (D-team).

Dit project 'Verbeteren en ontsluiten van het missiearchief Afghanistan' valt onder supervisie van [redacted], Hoofd Afdeling Evaluaties Defensiestaf (H-AEVAL). Het A- en D-team informeert periodiek H-AEVAL over de voortgang en ervaren knelpunten. Dit gebeurt middels mails en bijpraatsessies.

Activiteiten tot oktober 2022
Voor dit project zijn tenminste 508 gegevensdragers, variërend van losse harde schijven, zip drives en CD-ROMS, met informatie van de inzet in Afghanistan, beschikbaar bij het Semi-statisch Informatiebeheer (SIB) in Rijswijk. De werkzaamheden in het kader van het ordeningsproject vinden plaats in een TBB-2 ruimte op Camp New Amsterdam. Hier kan worden voorzien in toegang tot het HGI-netwerk, noodzakelijk gelet de classificatie van diverse datadragers en de daarop opgeslagen documenten. Het onderzoeksteam maakt voor ordening van de gegevens gebruik van '*IBM Watson Discovery'*, een geavanceerde ordenings- en zoekmachine die door gebruik te maken van *Artificial Intelligence* (AI) de doorzoekbaarheid van het dossier verbetert.

Zoektocht naar Special Report
Het project 'Verbeteren en ontsluiten van het missiearchief Afghanistan' kwam begin oktober 2022 in een stroomversnelling a.g.v. een NRC-artikel over een vermist rapport uit Afghanistan van 2010.[2] H-AEVAL verzocht het A-en D-team om te helpen met de zoektocht naar het, in het NRC-artikel, genoemde Speciale Report (SR). Vanaf dat moment hebben het A-team en D-team zich primair gericht op de zoektocht naar het SR. De teams hebben gedurende de periode medio oktober 2022 - eind februari 2023 ongeveer 1800 uur besteed aan dit 'ad hoc' onderzoek. Later zijn hier diverse deelonderzoeken bij gekomen.

Ten tijde van het verkrijgen van de aanvullende opdracht waren het A- en D-team bezig met het valideren van aannames om de, niet goed geordende en daardoor moeilijk vindbare, documenten in een geordende structuur te plaatsen. Daarvoor werd op dat moment gebruikgemaakt van een beperkte selectie uit de eerder genoemde 508 gegevensdragers. Deze selectie is gebaseerd op het aanbieden van een grote diversiteit aan bestandsformaten aan '*Watson'*, om zodoende de mogelijkheden van '*Watson'* te valideren en vervolgens door middel van AI de ordening van verdere data, te vergemakkelijken. Naar aanleiding van het verzoek voor het zoeken naar het SR is binnen de totaal beschikbare SIB-gegevensdragers vervolgens een aangepaste selectie gemaakt om zo gericht naar dit SR te kunnen zoeken. Gegeven de aangegeven periode waarin het SR mogelijk is geschreven

[1]Nota P-CDS, "Verbeteren en Ontsluiten Missiearchief Afghanistan,", nr BS2021014447 d.d. 01 juli 2021.
[2] Esther Rosenberg and Karel Berkhout, "De Gemartelde Boer, De Australische Special Forces En Een Verdwenen Rapport," *NRC*, 26 september 2022.

(april - juni 2010), zijn de gegevensdragers met het specifieke kenmerk “rotatie TFU-8”, aangezien deze rotatie deze tijdspanne omvat of gegevensdragers met data uit het jaar 2010, geselecteerd. Dit leverde uiteindelijk 12 SIB-gegevensdragers op waarop zich mogelijk het gezochte SR zou kunnen bevinden. Deze lijst met gegevensdragers is gevalideerd met de experts van het SIB en deze identificeerden dezelfde 12 gegevensdragers waarop in potentie het SR zou kunnen staan. De data van deze gegevensdragers is daarna versneld op het HGI-netwerk beschikbaar gemaakt. Op deze manier konden het A- en D-team met behulp van *‘Watson’* de beschikbare bestanden/ data doorzoeken. Het A-team heeft de resultaten ook nog, zeker bij twijfel over de indeling in de gemaakte dossierstructuur, handmatig (openen en inzien documenten) gecontroleerd.

De steller van het SR heeft aangegeven zijn document op het Humint Support System (HSS) te hebben opgeslagen. Mogelijke relevante gegevensdragers daarvan bleken in een kluis bij de DOPS/J2 te liggen. Dit was mogelijk het HSS gebruikt door eenheden van JISTARC in Afghanistan of een back-up hiervan. Dit bleek uiteindelijk te gaan om 42 gegevensdragers met daarop (mogelijk) de back-up van het HSS.

De installatie van ‘*IBM Watson Discovery’* verliep vanwege technische uitdagingen minder soepel dan verwacht. Daarom is in de zoektocht naar het Special Report tot nu toe (februari 2023) gewerkt met de iets minder geavanceerde versie van ‘*IBM Watson’*, te weten ‘*Explorer’*. Dat heeft geen invloed gehad op het (snel) kunnen doorzoeken van de aangeleverde gegevensdragers.

Omvang doorzochte data

Het totaal aantal documenten, ongeacht het type, van alle beschikbare datadragers op het HGI waar het Special Report mogelijk in zou kunnen staan, kwam daarmee op ~2.460.000 (4.0 TB). Van de in totaal 54 gegevensdragers, bleken 3 datadragers niet geheel leesbaar te zijn. Na interventie van een forensisch expert van JIVC konden alle relevante bestanden van de betreffende datadragers alsnog worden meegenomen in de zoektocht.

De inhoud van de 54 gegevensdragers is geïnventariseerd en op basis van een interview met de steller van het SR is besloten uitsluitend bestanden van het type Word (.doc, .docx) en PDF (.pdf) te selecteren. Hiervoor is gekozen omdat de steller aangaf dat hij het document sowieso in Word-format had opgesteld en opgeslagen. Het A-en D-team hebben, voor de volledigheid, besloten om ook PDF-bestanden in de selectie op te nemen. Dit omdat niet uitgesloten kon worden dat het Word-document ook als PDF-format opgeslagen kan zijn (of is ingescand). Er zijn ook 406 gecomprimeerde archieven (zip-bestanden) uitgepakt en de bestanden daaruit zijn opgenomen in de zoektocht. Van deze 406 gecomprimeerde archieven is uiteindelijk de inhoud van 404 meegenomen. Twee archieven waren corrupt en zijn hierdoor niet te openen. Na het toepassen van deze selectiecriteria bedroeg het aantal te doorzoeken documenten ~1.030.000.

Het proces van selectie, zowel gebruikt voor het project Missiearchief Afghanistan als voor de digitale zoektocht naar het SR, is grafisch weergegeven in een tweetal PowerPoint dia’s, opgenomen als bijlage A. In stap 4 en 5 wordt het aantal geselecteerde documenten van ~2.460.000 teruggebracht tot ~1.030.000.

De ~1.030.000 te doorzoeken documenten zijn vervolgens methodisch doorzocht met behulp van verscheidene, uiteenlopende zoektermen (zie bijlage B), combinaties en synoniemen van zoektermen, relaties van bestanden, betrokken personen en door het A-team met ervaring met en kennis van aanwezige documenten en de tijdsperiode.

In de methodisch doorzochte ~1.030.000 documenten (.doc, .docx en .pdf) zijn wel andere SR-bestanden gevonden, maar het gezochte SR is niet aangetroffen. Binnen deze methodische zoekactie

zijn 36 andere documenten gevonden waarvan de inhoud (deels) gerelateerd is aan het in het SR gerapporteerde incident (zie bijlage C). Deze informatie is deels verwerkt in de "Aangevulde NRC-tijdsbalk" (zie bijlage D). Dit om de beschikbare informatie in het relevante tijdsbestek zichtbaar te maken.

Bevindingen
Het A- en D-team hebben de zoektocht uitsluitend uitgevoerd op de aan hun ter beschikking gestelde gegevensdragers, afkomstig van SIB en uit de DOPS/J2-kluis (data van JISTARC).
Het A- en D-team concluderen dat:

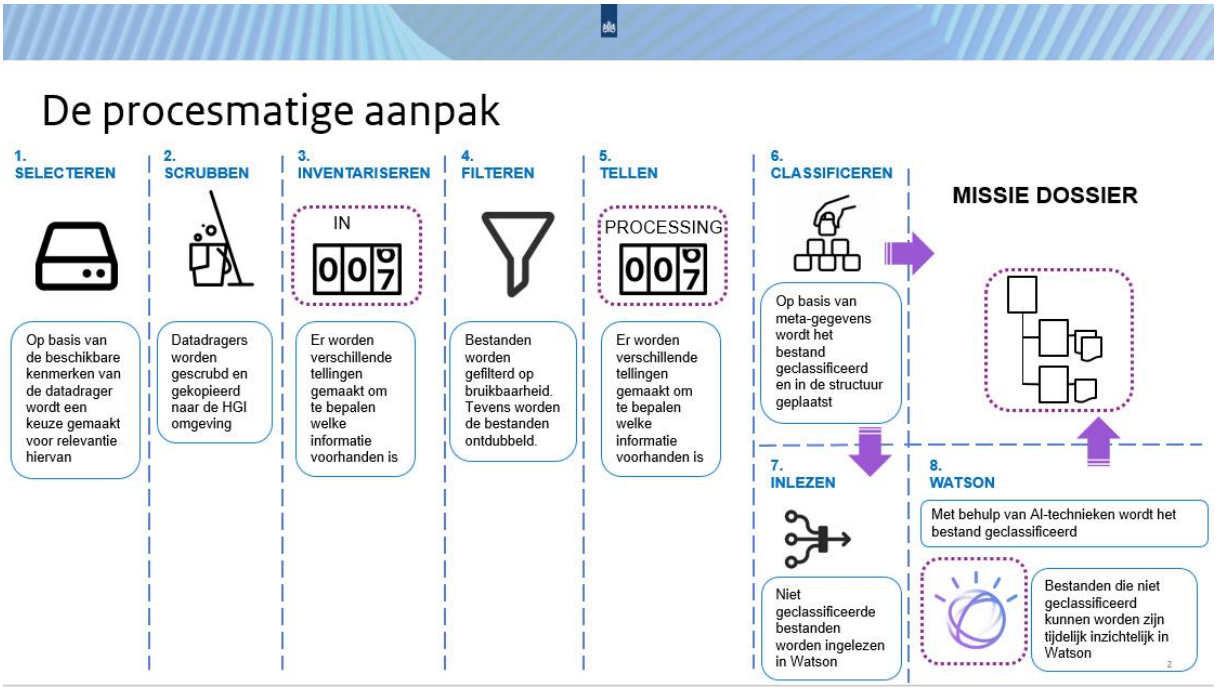
1. het gezochte SR is niet als .doc, .docx, of .pdf aangetroffen op de 12 van het SIB ontvangen datadragers;
2. het gezochte SR is niet als .doc, .docx of .pdf aangetroffen op de 42 van DOPS/J2 ontvangen datadragers;
3. ze geen uitspraak kunnen doen of op de DOPS/J2-datadragers het complete HSS staat. Daarover ontbreekt de kennis en dit zou gedaan moeten worden door een functionaris dit met dit systeem heeft gewerkt.

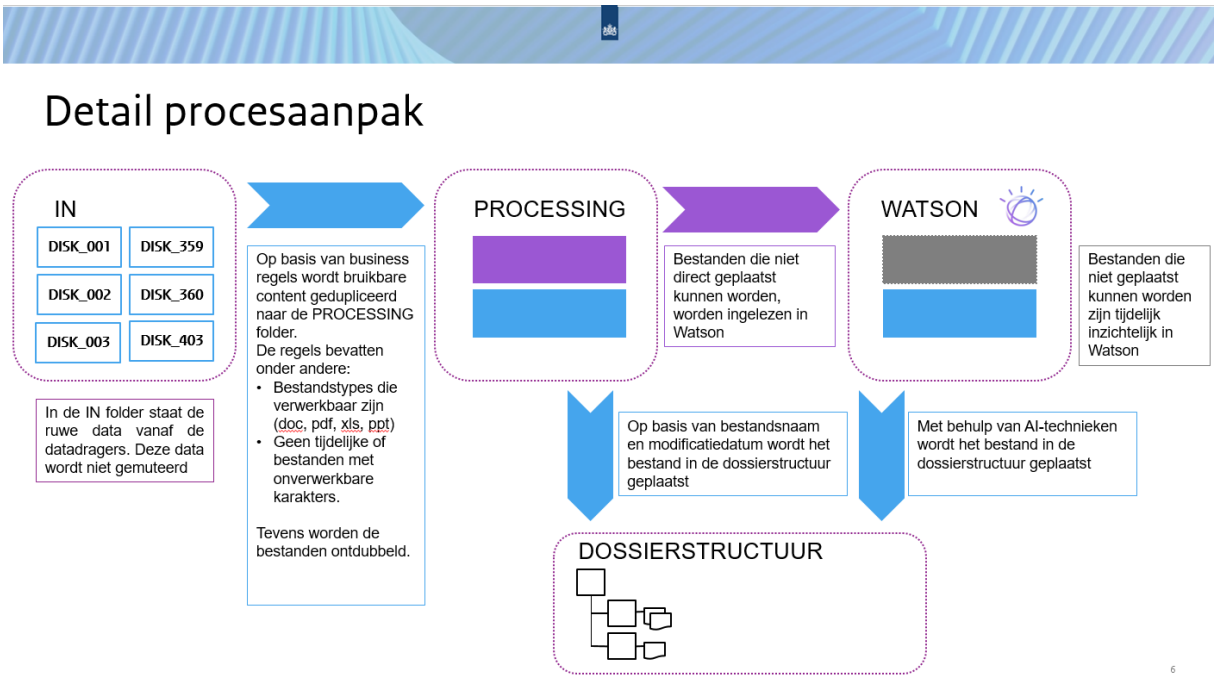
Personele capaciteit
A-team:

D-team:
3 *data scientists* van Accenture

Bijlage A	Procesmatige aanpak
Bijlage B	Zoektermen in relatie tot het SR
Bijlage C	Documenten die de NR- tijdlijn hebben aangevuld
Bijlage D	Aangevulde NRC-tijdlijn

Bijlage A: Procesmatige aanpak missie-archief Afghanistan





Bijlage B: Zoektermen in relatie het SR

De zoektermen die zijn gebruikt om het betreffende SR te vinden:[3]
10 mei/may
[redacted]
AFG-CV 0550
After action report /aar
After action review
Amb
Assessrep (ook uitgeschreven)
bayonet
bayonet [redacted]
BDAR
Casrep
Civcas
Coalition Forces
detainee (s)
[redacted]
Downrep
G2x
Gdsr
geweldmelding
[redacted]
grave
grave robbery
hammer
Humintrep (ook uitgeschreven)
incident report
Incspotrep
Intelligence report
Internal
Intimidation
Intsum (ook uitgeschreven)
[redacted]
Kia
killed
limited distribution
logboek
Logsitrep (ook uitgeschreven)
Misrep
Mrtf
[redacted]
Newsflash
NSI
NSI-TFH

[3] De zoektermen zijn ook in allerlei combinaties met elkaar toegepast, daarnaast zijn tevens alle bekende afkortingen en alternatieve schrijfwijzen van de zoektermen toegepast.

[redacted]
OM melding
[redacted]
order
owd (ook uitgeschreven)
Report civcas
[redacted]
[redacted]
Sigsum
Special report
specintrep (ook uitgeschreven)
stabbed
Summary
Supintrep
TF 66
Torture
Troops in contact /tic
West Dorafshan

Bijlage C: Relevante documenten, waaronder documenten die de NRC-tijdlijn hebben aangevuld

1. 17-04-2010 / Secret rel ISAF / Daily INTSUM 1355 TFU G2
2. 22-4-2010 / Secret rel ISAF / 211355 Dailyintsum
3. 24-04-2010 / Secret // rel AUS/CAN/BGR/USA/NLD / INTEL update 315
4. 25-04-2010 / Secret rel ISAF – not rel GCTF / (AUS) LNs in fear after Jalil's release
5. 26-04-2010 / STG Geheim / NLD eyes only / OWD
6. 27-04-2010 / NATO/ISAF Secret / TFU Humint Report 4276
7. 27-4-2010 / geen rubricering / SITREP CKM 0425
8. 28-04-2010 / Secret rel ISAF / INTSUM 1366 Joint Intel Cell TFU
9. 28-4-2010 / FST/G2X eyes only / interrogation report 1
10. 28-4-2010 / FST/G2X eyes only / interrogation report 2
11. 28-4-2010 / FST/G2X eyes only / interrogation report 3
12. 28-4-2010 / Secret rel ISAF / 1366 Daily intsum
13. 29-4-2010 / Secret rel ISAF / 1367 Daily intsum
14. 30-04-2010 / Secret//FGI AUS CAN GBR //rel to USA, ISAF, NATO / (AUS) Analyst notes for GRINTSUM
15. 30-4-2010 / NATO Secret rel ISAF / mullah jalil (obj Hammer)
16. 2-5-2010 / geen rubricering / sitrep CKM 0502
17. 3-5-2010 / Secret / monthly summary6-05-2010 / Secret//rel AUS/Can/GBR/NZL/USA/NLD / TF55 SUPINTREP 025
18. 7-5-2010 /Secret//rel AUS/CAN/GBR/USA/NLD / int-update 337
19. 7-05-2010 / Secret//ISAF/not rel GIROA / (AUS) mullah jalil sighted
20. 8-5-2010 / Secret/rel AUS/CAN/GBR/USA/NLD 025-ied strike-17april
21. 10-5-2010 / Secret rel ISAF / hammer
22. 11-05-2010 / Secret rel ISAF /) Mentoring Task Force INTSUM 1300
23. 12-5-2010 / geen rubricering / obj hammer KIA
24. 12-5-2010 /Secret rel ISAF / daily TK PRT 9
25. 12-5-2010 / Secret rel ISAF / 1380 dailyintsum
26. 15-5-2010 / Secret rel ISAF / meeting report 028 occp
27. 16-5-2010 / Secret rel ISAF / intsum 1304
28. 17-5-2010 / geen rubricering / sitrep CKM 0516
29. 21-5-2010 / NATO Secret rel ISAF / sprekers notities S2 0521
30. 24-5-2010 / geen rubricering / TAAA view on DRS-W
31. 27-5-2010 / Staatsgeheim Geheim rel ISAF NATO not GCTF or GIROA / veiligheidsupdate
32. 03-06-2010 / NATO Secret//rel ISAF / Daily INTSUM 3050 B-coy BG-12
33. 27-6-2010 / NATO Secret rel ISAF / gulam rasul
34. 1-7-2010 / geen rubricering / HHO 4
35. 4-7-2010 / geen rubricering / recent event list 13 marnscie
36. 18-7-2010 /NATO Secret rel ISAF / 1447 dailyintsum

Bijlage D: Aangevulde NRC-tijdlijn