



Belastingdienst

# Project Confiteor

Voortgang BPG en planning  
+ Extra agendapunt EGD

17-02-2023



# Extra agendapunt – Planning EGD

Selectie aan de Poort

Schending AVG

MSNP

Onderzoeken

Externe Gegevensdeling (EGD)

Bijzondere Persoonsgegevens (BPG)

Te lang intensief toezicht\*

Combineren

Geen EGD

EGD zonder tegemoetkoming

Geen BPG

BPG zonder tegemoetkoming

BPG met tegemoetkoming

Te lang intensief toezicht met tegemoetkoming

Informereren

TGVe

- Geen EGD of EGD zonder tegemoetkoming
- Geen BPG of BPG zonder tegemoetkoming

Erkennen

- Geen EGD of EGD zonder tegemoetkoming
- Geen BPG of BPG zonder tegemoetkoming of BPG met tegemoetkoming
- Te lang intensief toezicht met tegemoetkoming

\*Geen Te lang intensief toezicht wordt niet gecommuniceerd en Te lang intensief toezicht zonder tegemoetkoming is niet mogelijk.



# Planning EGD

December

Comm. plan  
voorleggen MT

Januari



Opstellen brief  
en flyer



Aanpassen  
bestaande  
belscrips

Februari



Testen brief  
en flyer



Aanpassen  
bestaande  
belscrips

24 Februari



Testen brief  
en flyer

28 Februari

Start  
communicatie

Start  
communicatie

Groepen B, D, F

Groepen C, E, G

## Uitdaging

- Eind 2022 eerste verzoek gedaan aan SSI voor effectbepaling EGD populatie en ook ontvangen.
- 31 januari een nieuwe update gevraagd aan SSI voor effecten EGD populatie.
- 9 februari is het verzoek doorgestuurd naar DOEN omdat een 'DOEN run' nodig is om effecten te bepalen voor briefverzending.

De data die we eerder vanuit SSI hebben ontvangen is niet geschikt voor briefverzending, resultaten eerst volgende 'DOEN run' voor EGD populatie zijn **9 maart** beschikbaar.



# Vaststellen notulen 03-02-2023

---

## Akkoord?



# Terugblik Actiepunten en besluiten 03-02-2023

	Omschrijving van besluit of actiepunt
<b>Besluit 1</b>	Als er Bijzondere Persoonsgegevens (BPG) in het vrije tekstveld (VTV) gevonden zijn, en het blijkt dat er ook nog bijlagen gekoppeld zijn aan de FSV registraties, worden de bijlagen niet meer onderzocht.
<b>Besluit 2</b>	De voorgestelde aanpak voor het onderzoeken van de FSV bijlagen is akkoord bevonden, met de kanttekening om een voorstel te maken voor een steekproef op de 'Geen BPG' populatie.
<b>Actiepunt 1</b>	Maak inzichtelijk hoeveel van de burgers waarvan BPG/SPG is vastgesteld in aanmerking komen voor tegemoetkoming.
<b>Actiepunt 2</b>	Stel een werkwijze op voor een steekproef op de bijlagen die voor medewerkers behapbaar en realistisch is. Dit zou eventueel ook ná verzending van brieven kunnen plaatsvinden.
<b>Actiepunt 3</b>	In kaart brengen welke bijlagen sowieso onderzocht moeten worden (burgers met effect). Hierbij ook de afspraak maken met SP om dit op te pakken.
<b>Actiepunt 4</b>	Als BPG geautomatiseerd uit de bijlagen gehaald kan worden dan zou dat goed zijn. Het niet-geautomatiseerde onderdeel kan wellicht met taal worden opgelost. Kom daarom vooral met voorstellen.
<b>Actiepunt 5</b>	In beeld krijgen per eind februari hoeveel mensen 375 euro en 750 euro gaan ontvangen zodat dit kan worden gecommuniceerd naar Finance en STAS ( <b>Actie Inzicht</b> )



# Inhoudsopgave

- 1 Aanpak analyse en zoektermen BPG / SPG
- 2 FSV Vrije tekstvelden
- 3 FSV Bijlagen
- 4 Planning
- 5 Bijlagen




# 1. Aanpak analyse en zoektermen BPG / SPG



# 1. Aanpak analyse en zoektermen BPG

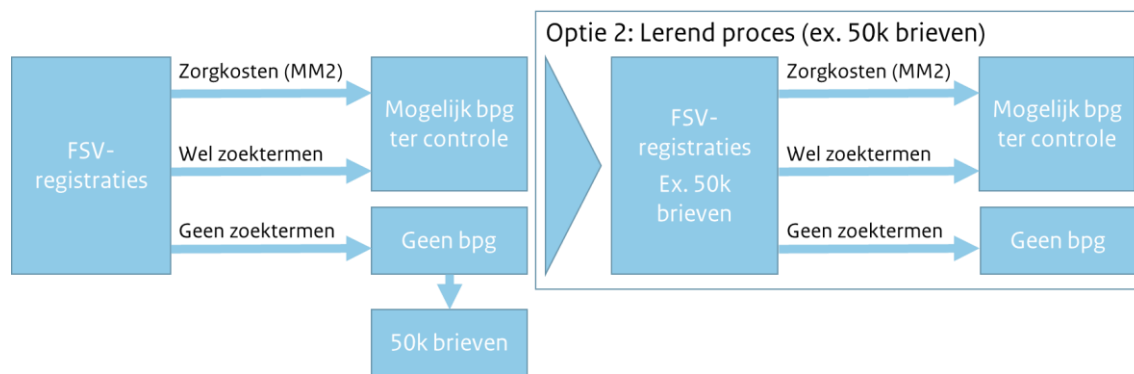
*De aanpak voor de analyse van Bijzondere persoonsgegevens (BPG) en de zoektermen zijn in augustus 2022 vastgesteld, daarbij is ook rekening gehouden met lerend vermogen.*

## Terugblik op eerder besluit

 In het MT van maandag **15-08-2022** is de aanpak voor het vaststellen van de zoektermen voor het BPG onderzoek vastgesteld. Er is een initieel document met zoektermen voorgesteld met twee opties m.b.t. de aanpak:


**Optie 1:** de zoektermen zijn definitief en we vullen deze niet meer aan met nieuwe inzichten

**Optie 2:** we implementeren toekomstige inzichten wél in zoektermen, maar passen deze niet meer toe op de 50k populatie



**Het MT heeft daarna het volgende besluit genomen:** MT niet eens met één van beide voorgestelde opties, echter opteert MT een optie 3: Waarbij de zoektermen in eerste instantie worden vastgesteld op het moment dat we minstens 95% zeker denken te zijn dat we de juiste brief versturen. Daarnaast blijven we leren en indien we niet volledig zijn geweest, dan dienen we een brief na te sturen.

## Vervolg

 **Actie na MT:** Voorstel maken over hoeveel leerslagen/iteratieslagen dienen te worden genomen over de set aan zoektermen om op een foutmarge **van Max5% onjuiste brieven** te komen, die later dienen te worden herzien. M.a.w.. hoe komen we tot een set aan zoektermen waarmee we voor minstens **95% zeker** zijn dat de juiste brief wordt gestuurd.

- De zoektermen zoals gepresenteerd in het MT van 15-08-2022 zijn aangevuld met alternatieve schrijfwijzen en spellingsfouten en op **28-08-2022** is deze lijst in gebruik genomen. Deze lijst met zoektermen heeft lang gezorgd voor 95% zekerheid. De werkwijze voor de analyse van Vrije tekstvelden is als volgt:
  - We gaan niet automatisch concluderen dat een registratie BPG bevat -> **alle zoektermen leiden tot handmatige controles.**
  - De registraties met een verwijzing naar zorgkosten uit MM2 beschouwen we direct als mogelijke BPG.
- In de loop van de tijd zijn er meer registraties van verschillende tabbladen ontvangen (bijv. Tips & Kliks) en toen is gebleken dat we met de zoektermen de 95% zekerheid niet meer halen. Daarom is de lijst aangevuld (er zijn dus **geen zoektermen verwijderd!**) met **55 woorden** (incl. verschillende schrijfwijzen).
- Voor het onderzoek naar de bijlagen is een zoektermenlijst opgesteld om documenten die vallen onder de Wet Politiegegevens te identificeren.





# Bestluitvorming | Goedkeuring herijkte zoektermen

## GEWENSTE BESLUITVORMING

Gaat het MT akkoord met het voorstel om:

- De lijst met zoektermen (inclusief de 55 toegevoegde woorden) vast te stellen voor de definitie van 'mogelijk BPG' voor verder onderzoek.

## NADERE TOELICHTING

### SITUATIE

In augustus 2022 is het onderzoek naar BPG in FSV gestart met een zoektermenlijst zoals in het MT van 15-08-2022 besproken. In de loop van de tijd is deze lijst aangevuld met zoektermen omdat de minimale 95% zekerheid niet werd behaald vanwege nieuwe registraties op tabbladen die niet eerder zijn onderzocht. Om de minimale 95% zekerheid te halen zijn er 55 woorden (incl. alternatieve schrijfwijzen) in de afgelopen maanden toegevoegd aan de lijst om te voldoen aan de norm van minstens 95% zekerheid.

### VOORSTEL

De zoektermenlijst is een middel om de mogelijkheid van BPG in FSV registraties te identificeren en dit leidt niet tot een geautomatiseerde conclusie want alle registraties met een hit op de zoektermen worden handmatig door het SP gecontroleerd. Het aanvullen van deze zoektermenlijst om tot minstens 95% zekerheid te komen is onderdeel van het lerend vermogen. Daarom is het voorstel om de zoektermenlijst zoals gepresenteerd in de bijlage '20230215\_BPG-SPG zoektermenlijst' goed te keuren.



# Bestluitvorming | Goedkeuring WPG zoektermen

## GEWENSTE BESLUITVORMING

Gaat het MT akkoord met het voorstel om:

- De zoektermen voor het vaststellen van documenten die vallen onder de Wet Politiegegevens (WPG) goed te keuren.

## NADERE TOELICHTING

### SITUATIE

Documenten en informatie die onder de Wet politiegegevens (Wpg) valt, mag alleen ingezien worden door een Buitengewoon Opsporingsambtenaar (BOA). Uit eerdere analyses is bekend dat (met name in de T&K en IVZ registraties, maar ook AFS en POS registraties) dergelijke bijlagen aanwezig zijn. Het is daarom zaak om deze WPG documenten (zo veel mogelijk) zonder tussenkomst van BZB medewerkers te identificeren en niet te lezen/openen. In het MT overleg van 26-01-2022 heeft het MT goedgekeurd dat de aanwezigheid van bijlagen waarin zoektermen zoals “verhoor”, “procesverbaal”, “PV [spatie]”, “politie” etc. voorkomen, moet leiden tot de conclusie dat deze documenten vallen onder de WPG en er dus sprake is van SPG zonder de documenten verder te bekijken.

### VOORSTEL

In lijn met de eerder verleende goedkeuring vragen we aan het MT om de WPG zoektermenlijst zoals gepresenteerd in de bijlage ‘20230215\_BPG-SPG zoektermenlijst’ goed te keuren.

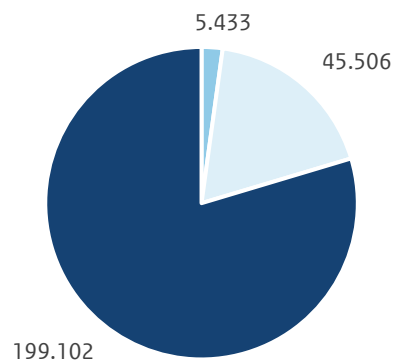


## 2. FSV Vrije tekstvelden (VTV)



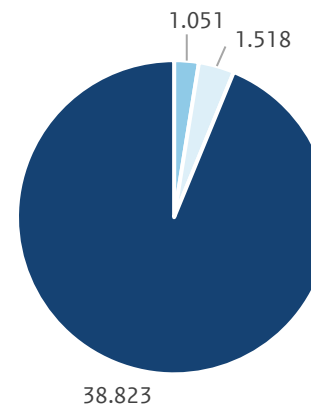
## 2. FSV Vrije tekstvelden (VTV) – Stand van zaken 17-02-2023

Aantal BSN's (250.041) waarvan het vrije tekstveld is onderzocht op BPG



■ Te onderzoeken BSN's ■ BSN's met BPG in vrij tekstveld ■ BSN's zonder BPG in vrij tekstveld

Aantal BSN's met bijlagen (41.392) waarvan het vrije tekstveld is onderzocht



■ Te onderzoeken BSN's ■ Wel BPG ■ Geen BPG

### Status

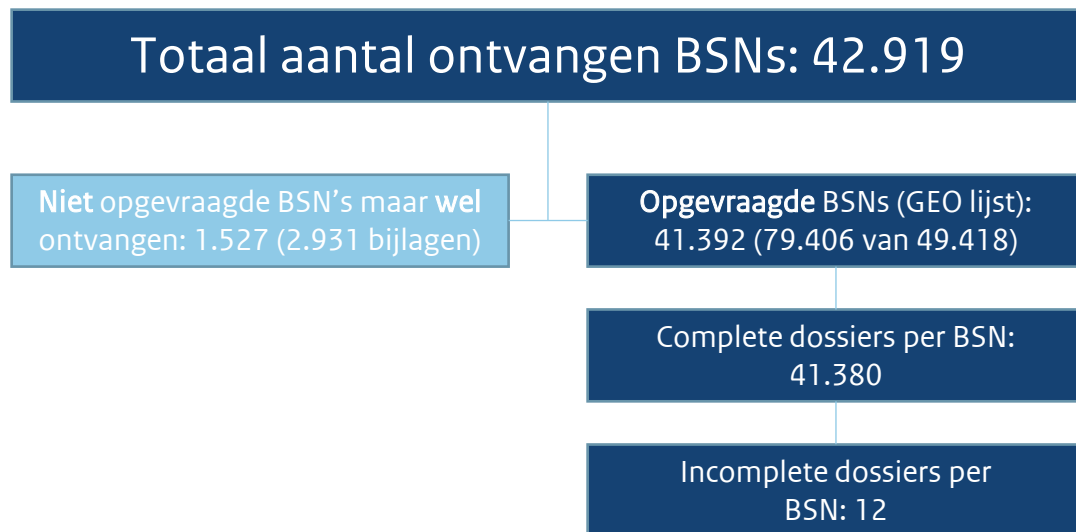
- Van in totaal **6.484 burgers** worden de vrije tekstvelden nog onderzocht, dit zijn **overledenen** en deze laatste batch staat nog uit bij SP.
- Van de 47.024 burgers met BPG in de vrije tekstvelden komen **circa 7.500 burgers** in aanmerking voor een **potentiële tegemoetkoming**.



## 3. FSV Bijlagen



## 3.1 FSV Bijlagen – Overzicht ontvangen bijlagen



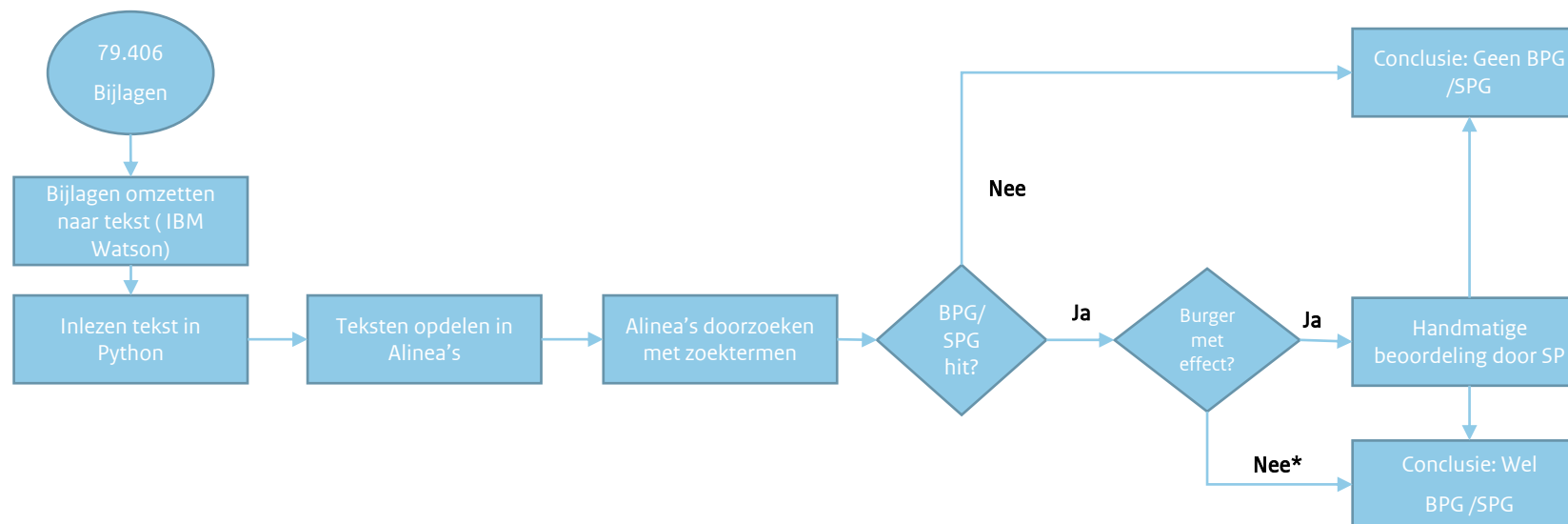
BSN	Aantal bijlagen verwacht	Aantal bijlagen gekregen	Vershil
BSN #1	10	9	-1
BSN #1	3	2	-1
BSN #3	4	3	-1
BSN #4	4	3	-1
BSN #5	13	12	-1
BSN #6	3	2	-1
BSN #7	14	13	-1
BSN #8	5	4	-1
BSN #9	1	0	-1
BSN #10	1	0	-1
BSN #11	1	0	-1
BSN #12	1	0	-1
<b>Totaal</b>	<b>60</b>	<b>48</b>	<b>12</b>

### Status

- Voor BPG zijn na collectie de bestanden uniek hernoemd en tekstueel doorzoekbaar gemaakt met Watson.
- De bijlagen voor 12 BSN's zijn nog niet compleet, waarvan we bij 4 BSN's überhaupt geen bijlage hebben ontvangen. Hiervoor moeten de bestanden handmatig uit FSV worden geëxporteerd, dit proces loopt.



## 3.2 FSV Bijlagen – Plan van aanpak en resultaten



\* De zoektermen worden nog geanalyseerd om te bepalen welke zoektermen direct gekoppeld kunnen worden aan een categorie.

Bij het onderzoek in de vrije tekstvelden is een ruim scala aan zoektermen gebruikt omdat er een handmatige analyse plaatsvond.

Daardoor kunnen we ook met grote zekerheid vaststellen dat er geen BPG / SPG in de bijlagen voorkomt. **Het voorstel is daarom ook om achteraf nog een steekproef uit te voeren en niet meer voor verzending.**

Resultaten		
Categorie	Aantal BSN	Vervolg
SPG	9.927	<b>Afgerond</b>
SPG + BPG	6.464	Besluit nodig
BPG	15.612	Besluit nodig
Geen SPG en BPG	5.047	<b>Afgerond</b>
BPG in VTV	4.342	<b>Afgerond</b>
<b>Totaal</b>	<b>41.392</b>	

### Te onderzoeken BSN's conform bovenstaande aanpak

- **1.698 unieke burgers** hebben een bijlage en krijgen **potentieel tegemoetkoming**
- Voor **959** van de 1.698 unieke burgers is er in **de VTV al BPG/SPG gevonden** en/of zijn er **WPG** zoektermen gevonden. Zij krijgen een potentiële tegemoetkoming.
- Voor de rest (**739 burgers**) van de 1.698 unieke burgers moet er **gecontroleerd** worden of er BPG/SPG is.
- Voor **601** van de 739 burgers is er **mogelijk BPG/SPG** in de naam van het bestand en/of de inhoud van het bestand of zijn niet alle bijlagen automatisch in te lezen.



# Bestluitvorming | Geautomatiseerd BPG vaststellen

## GEWENSTE BESLUITVORMING

Gaat het MT akkoord met het voorstel om:

- Optie 1: Bijlagen met een hit op de BPG zoektermen waar mogelijk geautomatiseerd vaststellen, overige bijlagen niet onderzoeken en binair BPG communiceren .
- Optie 2: Bijlagen met een hit op de BPG zoektermen waar mogelijk geautomatiseerd vaststellen, overige bijlagen handmatig onderzoeken.

## NADERE TOELICHTING

### SITUATIE

Bij het onderzoek naar BPG in de vrije tekstvelden zijn alle registraties met een hit onderzocht door het SP, daardoor zijn er een aantal zoekwoorden die niet automatisch aan een BPG categorie te koppelen zijn. Zo is een hit op 'Algerijn' te koppelen aan nationaliteit, maar de zoekterm 'Alge' is dat niet omdat Algerijnse of algemeen kan zijn. Door gebruik van de zoekterm 'Alge' vergroot je de zekerheid maar dat kan alleen als er een handmatige controle volgt. Daarom zouden we op basis van de zoekwoorden een deel geautomatiseerd aan een categorie kunnen koppelen en het overige deel binair (wel/geen BPG) kunnen communiceren of handmatig de categorie door SP kunnen bepalen.

### VOORSTEL

**Optie 2: Bijlagen met een hit op de BPG zoektermen waar mogelijk geautomatiseerd vaststellen, overige bijlagen handmatig onderzoeken.**

**Met de aanvulling dat we dit doen voor burgers waarvoor geen enkele categorie geautomatiseerd gevonden is.**





# Bestluitvorming | Onderzoek BPG als SPG is vastgesteld

## GEWENSTE BESLUITVORMING

Gaat het MT akkoord met het voorstel om:

- Optie 1: Wanneer er SPG is vastgesteld in één van de bijlagen, en BPG in een andere bijlage, dan onderzoeken we de bijlage met BPG alleen als de burger een effect heeft.
- Optie 2: Wanneer er SPG is vastgesteld in één van de bijlagen, en BPG in een andere bijlage, dan onderzoeken we de bijlage met BPG niet.

## NADERE TOELICHTING

### SITUATIE

Er zijn burgers met meerdere bijlagen in FSV, het komt dus voor dat een burger in een bijlage heeft dat geclassificeerd is als WPG document en een andere bijlage met een hit op BPG. Op basis van het WPG document wordt SPG vastgesteld voor de burger, de vraag is dan of het nog noodzakelijk om de overige bijlagen te beoordelen op eventuele BPG als er een hit is op de zoektermen. De bijlagen van burgers met effecten worden in ieder geval, conform eerder besluit, handmatig beoordeeld door het SP.

### VOORSTEL

Als er voor optie 2 is gekozen bij het vorige besluit, hetzelfde besluit ook voor deze categorie volgen.

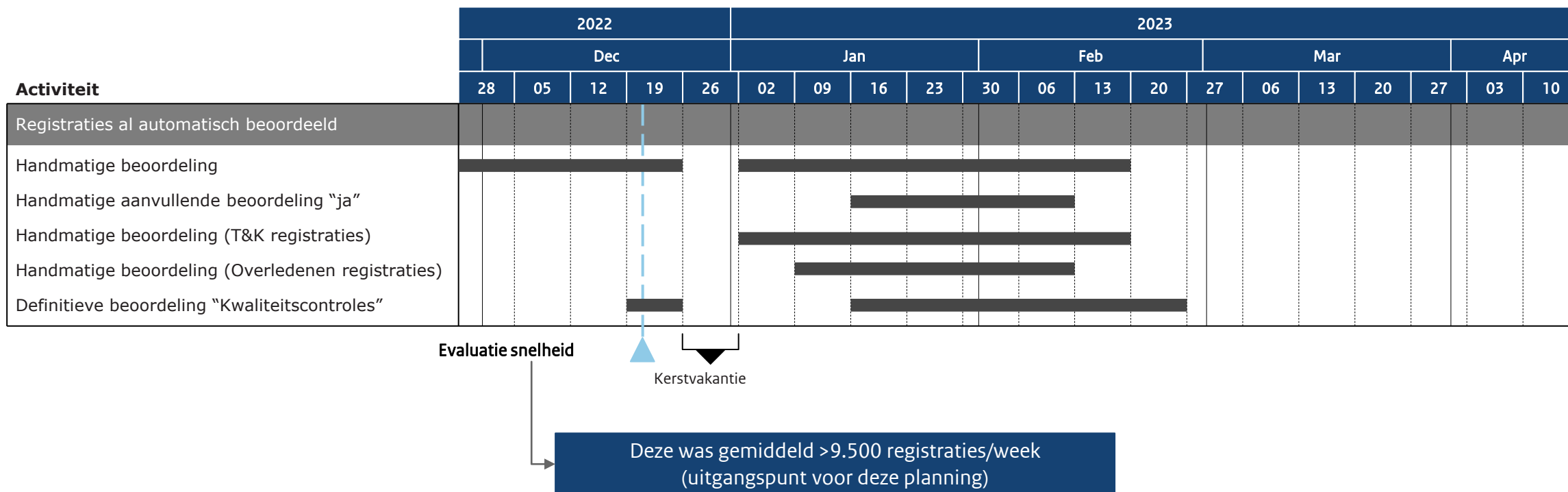


# 4. Planning



## 4. Planning – Vrije tekstvelden (VTV)

*De planning is dat rond medio februari de analyse van de vrije tekstvelden gereed is. Kwaliteitscontroles op de resultaten van de VTV analyse zijn eind februari afgerond.*












## 5. Bijlagen





# 5.1 Bijlage – Context BPG

*BPG/SPG analyseren we in het kader van de opvolging van motie Snels. We onderzoeken registraties uit de FSV met en zonder bijlagen.*

## Context

-  BZB onderzoekt welke FSV-registraties **bijzondere of strafrechtelijke persoonsgegevens bevatten (BPG/SPG)** in het kader van de compensatieregeling voor burgers in de FSV en het creëren van transparantie;
-  Om dit uit te kunnen voeren moeten er specifiek BPG/SPG geanalyseerd worden. We doen dit in het kader van een zwaarwegend algemeen belang zodat de burgers **gecompenseerd** kunnen worden voor het feit dat hun BSGP in de FSV geregistreerd stonden;
-  Dit is in lijn met de **motie Snels** die de opdracht geeft om een compensatieregeling te ontwikkelen en de berichtgeving aan de Tweede Kamer omtrent de vormgeving van de tegemoetkoming voor de FSV;
-  Het BPG/SPG onderzoek richt zich specifiek in het algemeen op de **vrije tekstvelden** en de **bijlagen** in de FSV;
-  Uit **eerdere onderzoeken** bleken op deze locaties bijzondere en/ of strafrechtelijke persoonsgegevens te kunnen staan.

## Doelstelling en scope

-  De doelstelling van het analyseren van bijlagen is tweeledig:
  - Het identificeren van de aanwezigheid van bijzondere persoonsgegevens (BPG) in bijlagen en categoriseren naar soort BPG;**
  - Het identificeren van de aanwezigheid van strafrechtelijke persoonsgegevens (SPG) in bijlagen;**
-  De scope bevat de bijlagen behorende bij:
  - Alle BSNs uit de FSV (ongeacht effect) voor de aanwezigheid van BPG;**
  - BSNs met een kwalificerend effect voor compensatie voor de aanwezigheid van SPG (zie randvoorwaarden en uitgangspunten, volgende pagina);**
  - Bijlagen die niet onder de Wet politiegegevens e.d. vallen waarvoor inzage vereist dat de lezer/verwerker een buitengewoon opsporingsambtenaar (BOA) is;**



## 5.2 Specifieke aanpak Wet Politiegegevens documenten

*Documenten die alleen door BOAs verwerkt mogen worden identificeren we eerst op basis van de bestandsnaam en daarna op op basis van de inhoud*

Documenten en informatie die onder de Wet politiegegevens (**Wpg**) valt, mag alleen ingezien worden door een Buitengewoon Opsporingsambtenaar (**BOA**). Uit eerdere analyses is bekend dat (met name in de T&K en IVZ registraties, maar ook AFS en POS registratieis) dergelijke bijlagen aanwezig zijn. Het is daarom zaak om deze Wpg documenten (zo veel mogelijk) zonder tussenkomst van BZB medewerkers te identificeren en niet te lezen/openen.

Dergelijk documenten kunnen volgens onderstaand stappen plan mogelijk geïdentificeerd worden.

1. We controleren alleen SPG van burgers die in aanmerking komen voor tegemoetkoming
2. Analyse van de documentsnaam. Door deze te doorzoeken op het voorkomen van woorden zoals “verhoor”, “procesverbaal”, “PV [spatie]” en “politie”. Kunnen deze documenten al mogelijk geïdentificeerd worden.
3. Vervolgens kan op de inhoud van documenten ook met een dergelijke lijst gezocht worden om dergelijke documenten te identificeren.
4. Additioneel is Watson (door eerdere herkenning) in staat een dergelijk document te identificeren.

### **Besluit**

De aanwezigheid van dergelijke bijlagen, waarin de zoektermen voorkomen, moet leiden tot de conclusie dat er sprake is van SPG zonder verder te bekijken.





Belastingdienst