

Personenportal WO2

The screenshot shows a web browser window with the URL `dati.cdec.it/odlive/?http://dati.cdec.it/od/shoah/person/4919`. The main content is a semantic network diagram with nodes and relationships:

- Labi, Grazia** (red circle) is connected to **Labi, Isacco** (red circle) via the relationship `« consortOf »`.
- Labi, Grazia** is connected to **Labi, Rubina** (red circle) via the relationship `joint`.
- Labi, Isacco** is connected to **Labi, Rubina** via the relationship `« parentOf »`.
- Labi, Rubina** is connected to **persecution of Labi, Rubina 25/9/1934 Tripoli** (purple circle) via the relationship `persecution »`.
- persecution of Labi, Rubina 25/9/1934 Tripoli** is connected to **liberazione di Labi, Rubina** (teal circle) via the relationship `liberation »`.

On the right side, there is a metadata panel for the selected entity (Labi, Rubina):

| | |
|--------------------|---|
| type | Person |
| label | Labi, Rubina |
| gender | F |
| first name | Rubina |
| family name | Labi |
| source | Il libro della memoria : gli ebrei deportati dall'Italia, 1943-1945 / Liliana Picciotto ; ricerca della Fondazione Centro di documentazione ebraica contemporanea. - Ed. 2002: altri nomi ritrovati. - Milano : Mursia, 2002, pp. 77-80, pp. 66-71. |
| provenance | db deportati |

At the bottom left, there is a search bar with the URL `http://lod.xdams.org/sparql` and a list of search results:

- Labi, Rubina
- `http://dati.cdec.it/od/shoah/person/4950`

Waarom een Personenportal?

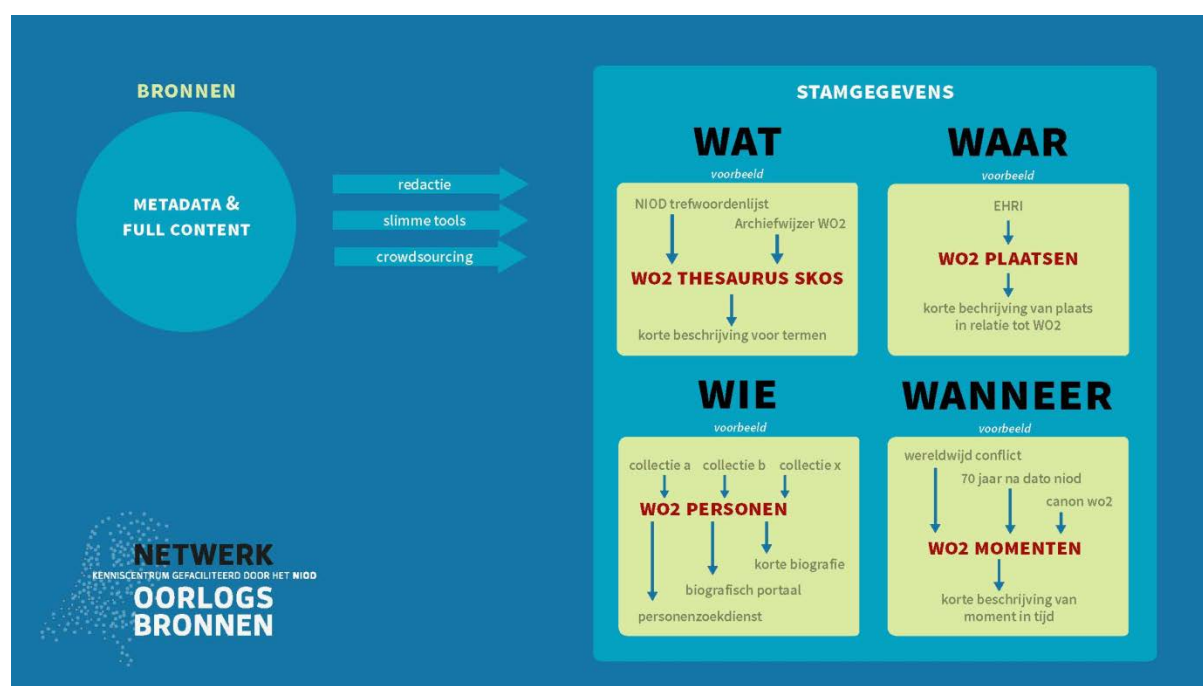
Achtergrond

Het Netwerk Oorlogsbronnen is opgericht om WO2 gerelateerde collecties, verspreid over 400 organisaties in Nederland en even zoveel meer in het buitenland, digitaal samen te brengen en context te geven. Dit laatste is belangrijk omdat veel digitale bronnen, zodra ze gedigitaliseerd worden en in een nieuwe online context geplaatst worden, de relaties met hun oorspronkelijke context kwijtraken.

De kleertjes van Jaap Wertheim, die hij droeg toen hij in onderduik ging, kunnen bijvoorbeeld op een portalsite hun emotionele lading verliezen en als een schattig groen kinderpakje beschouwd worden.

Het Netwerk Oorlogsbronnen heeft inmiddels bijna 10 miljoen bronnen samengebracht: variërend van krantenartikelen en archiefbeschrijvingen tot foto's en museale objecten. De komende jaren komen daar nog miljoenen digitale objecten bij.

Vanuit zijn taak om de verzamelde digitale WO2-collecties te contextualiseren is het Netwerk Oorlogsbronnen in 2016 gestart met de ontwikkeling van digitale instrumenten rond de wie-wat-waar-wanneer assen.



In bovenstaande afbeelding uit 2015 is het contextualiseren van de bronnen van Oorlogsbronnen schematisch weergegeven.

Een overzicht van de digitale instrumenten in ontwikkeling:

De Thesaurus_wo2 (https://niod.data.nl/Thesaurus_wo2.html) is voor de WAT-as ontwikkeld.

De WAAR-as: In de thesaurus zijn inmiddels relevante plaatsen, inclusief geocodes (longitude, latitude, in Nederlands-Indië wellicht ook altitude), opgenomen. In 2017 worden alle "named places" in de objecten van Oorlogsbronnen van geocodes voorzien. De techniek hiervoor wordt op dit moment uitgerold.

Voor de WANNEER-as is in samenwerking met eHumanities/VU/UVA de Events-Workinggroup opgericht. Naast het structureren van in bronnen genoemde tijdstippen en het toevoegen van belangrijke data (D-Day, Dolle Dinsdag enzovoorts) aan de WO2 thesaurus streeft het NOB ernaar om micro-gebeurtenissen (onderduik van Jaap Wertheim) te koppelen aan macro-gebeurtenissen (Jodenvervolging, WO2).

De Events-workinggroup werkt op dit moment fondsen om in kranten en egodocumenten gebeurtenissen (op microniveau) te detecteren, te classificeren en te verbinden met events op macroniveau. Doel is om de computer te trainen dit werk te doen.

De Personenportal WO2, die in dit projectplan uitgewerkt staat, start in 2017 en geeft invulling aan de WIE-as

Samenwerking

Doel van het project Personenportal is om persoonsgegevens op gestructureerde wijze samen te brengen, persoonsgegevens van verschillende organisaties met elkaar te vergelijken en door zo veel mogelijk materiaal te (trachten te) matchen wil het Netwerk Oorlogsbronnen een register van Personen aanleggen.

Per persoon wordt inzichtelijk wie het was en waar gegevens over de betreffende persoon te vinden zijn.

Voorbeeld:

Joseph Marie van de Mortel (Tilburg, 5-4-1919 –Vught, 10-8-1944)

<https://oorlogsgravenstichting.nl/persoon/106424/joseph-marie-van-de-mortel>

is:

<http://www.nmkampvught.nl/biografieen/mortel-joost-van-de/>

- o is zoon van:

<http://www.regionaalarchief Tilburg.nl/zoek-eeen-foto/detail/79234f26-85d0-47b2-a50a-a6c769691f32/media/c1e41df4-3285-4cc5-acd3-7cd10a9b8873?mode=detail&view=horizontal&q=Joseph%20Marie%20van%20de%20Mortel&rows=1&page=4>

4

- o Hij wordt genoemd op de erelijst: <http://www.erelijst.nl/>

- o Het NIOD beschikt over een Kamp Vught kaart van hem.

- o Het Rode Kruis beschikt over een dossier.

- o Het CBG beschikt over bronnen.

- o In de databank van Oorlogsbronnen bevinden zich foto's van hem.

<http://www.regionaalarchief Tilburg.nl/zoek-eeen-foto/detail/4f4bf61d-7131-45c5-927f-8968aef366ca/media/58487c7a-dcd1-4e35-bcfe-7d05830ee733?mode=detail&view=horizontal&q=Joseph%20Marie%20van%20de%20Mortel&rows=1&page=2>

2

- o Zijn naam wordt genoemd op de Fusillade Plaats, een monument vlakbij Kamp Vught.:

<http://www.4en5mei.nl/herdenken-en->

[ieren/oorlogsmonumenten/monumenten_zoeken/oorlogsmonument/2074](http://www.4en5mei.nl/herdenken-en-ieren/oorlogsmonumenten/monumenten_zoeken/oorlogsmonument/2074)

Het doel van dit project is echter niet om alleen een nieuw register van personen aan te leggen: we willen personen verbinden aan plaatsen, gebeurtenissen en bronnen. Hierdoor wordt ook inzichtelijk dat J.M. van de Mortel, zoon van de Burgemeester van Tilburg, lid was van De Trouw-groep. Wat De Trouw-groep was. Wie er nog meer lid van waren, enzovoorts.

Bovendien wordt inzichtelijk in welke bronnen en collecties informatie te vinden is over J.M. van de Mortel.

Door samen te werken en gegevens te *matchen* met een heel brede groep bron behorende instellingen hopen we op het niveau van individuen toegang te creëren tot zo veel mogelijk relevante gegevens.

In dit project werken de volgende organisaties samen om gegevens over personen beschikbaar te maken:

| |
|------------------------------------|
| Organisatie |
| Centraal Bureau voor de Genealogie |
| HC Kamp Westerbork |
| Joods Cultureel Kwartier |
| Nationaal Archief |
| Nationaal Comité 4 en 5 mei |

| |
|-----------------------------------|
| Ned. Rode Kruis/Oorlogsarchief |
| Netwerk Oorlogsbronnen |
| NIOD |
| NM Kamp Vught |
| NM Kamp Amersfoort |
| Oorlogsgravenstichting |
| Oranjehotel |
| St. Vriendenkring Neuengamme etc. |

Het Netwerk Oorlogsbronnen is op dit moment druk bezig nieuwe collecties aan te sluiten en brengt in dit project in ieder geval de volgende collecties in:

| |
|--|
| Collecties |
| Delpher Kranten, oorlogskranten |
| Beeldbank WO2 |
| Archieven WO2 |
| 8 Geheugen van Nederland collecties |
| Collecties Erfgoed van de Oorlog |
| NIOD archieven |
| EHRI archieven |
| Beeldbanken van het Historisch Centrum Overijssel, Digitaal Limburg en de gemeentes Den Haag, Leiden, Wageningen, Tilburg, Weert, Nijmegen en relevante collecties uit de beeldbank van het Nationaal Archief. |
| Museale collecties uit de Collectie Nederland (voorheen DIMCON). |
| Objecten WOII van de Atlas van Stolk, Amsterdam Museum, Museon, Rotterdam Museum en het Rijksmuseum. |

Nader ingezoomd op de personenportal

De personenportal is voor het Netwerk Oorlogsbronnen de digitale as waar individuele mensen te vinden zijn. Doel van de WIE-as is om mensen in bronnen te kunnen herkennen en aan te kunnen geven dat één individu in verschillende bronnen voorkomt.

De personenportal is een webwijzer. Het is een online zoekmachine ingericht om bronnen over specifieke personen te vinden. Over de persoon zelf zal beperkte informatie te vinden zijn om het bronnenmateriaal te kunnen duiden en vast te kunnen stellen of het de juiste persoon betreft. De personenportal is geen nieuwe biografische databank zoals het Joods Monument en de website van de Oorlogsgravenstichting al gerealiseerd hebben. Deze websites zijn al heel erg goed en uitgebreid in het vinden, onderhouden en aanvullen van biografische gegevens. De personenportal zal als webwijzer graag gebruik maken van basisgegevens uit deze portals en mensen voor biografische informatie doorverwijzen.

Het liefst linken we vanuit de personenportal direct door naar de gedigitaliseerde bronnen over de betreffende persoon. Maar als techniek, wet- en regelgeving of beleidskeuzes dit het onmogelijk maken, verwijst de personenportal door naar de instelling waar het bronnenmateriaal beschikbaar is.

De personenportal beslaat het hele spectrum van de Tweede Wereldoorlog: Joodse slachtoffers, militairen, slachtoffers in Nederlands-Indië, te werk gestelden, gevangenen, verzetsstrijders, collaborateurs, SS-vrijwilligers enzovoorts. Het Netwerk Oorlogsbronnen heeft de ambitie om iedereen in de personenportal op te nemen die in bronnen met betrekking tot de Tweede Wereldoorlog voorkomt.

Ook het begrip bronnen is breed gedefinieerd:

De bronnen waarmee we de personenportal op willen bouwen zijn formeel gezien afgeleide bestanden van oorspronkelijk bronnenmateriaal met persoonsgegevens. Deze bestanden (namenindexen) vormen de basis van de personenportal.

Aan de hand van deze namenindexen willen we primair zichtbaar maken wie in welke archiefstukken, persoonskaartenbanken of andere bronnen zoals krantenartikelen (uit de oorlog) of beeldbanken genoemd worden.

In tweede instantie streven we er ook naar om secundaire bronnen zoals boeken, tijdschriftartikelen, websites (dus bronnen die na de oorlog geproduceerd zijn) over de betreffende persoon toegankelijk te maken.

Doelgroepen

De Personenportal wordt primair opgebouwd om:

- Onderzoekers
- Directbetrokkenen/nabestaanden
- Medewerkers van de betrokken partner instellingen

te helpen om relevante gegevens over individuen te vinden.

De personenportal wordt een zoekingang op namen, waarna de onderzoeker een lijst met mogelijkheden krijgt. Vervolgens, als hij/zij een keuze maakt wordt beperkte biografische informatie en een bronnenlijst gepresenteerd.

Sommige bronnen zijn direct digitaal raadpleegbaar, wellicht zijn sommige bronnen voor partners via een login module raadpleegbaar en in andere gevallen wordt alleen gemeld dat bepaalde instellingen over relevante bronnen beschikken.

De beoogde personenportal wordt niet gebouwd als herdenkingsplaats, zoals het Joods Monument en de website van de Oorlogsgravenstichting. We zetten niet in op het ontwikkelen van publieksdiensten of interfaces, maar vooral op het samenbrengen en toegankelijk maken van bronnen.

Doel van de personenportal van Netwerk Oorlogsbronnen is om mensen naar bronnen en bron beheerende instellingen te verwijzen als ze op zoek zijn naar een specifieke persoon. Dit zullen vooral onderzoekers en genealogen/familieleden zijn. Wij hopen hier ook medewerkers van instellingen mee te helpen, die vaak onderzoek doen naar kampslachtoffers, verzetsstrijders enzovoorts en met veel verschillende organisaties contact op moeten nemen.

De personenportal ontwikkelt zelf geen eigen biografische gegevens¹: de persoonsgegevens worden uit de systemen van de partners opgehaald en gematcht met andere namenlijsten. Met behulp van scripts wordt berekend hoe waarschijnlijk twee gelijkende namen uit twee bronnen betrekking hebben op dezelfde persoon. Er wordt een Identifier toegekend en daarna wordt geanalyseerd welke (secundaire-) bronnen matchen. Deze verrijkte persoonsbeschrijvingen zijn raadpleegbaar in de personenportal en kunnen (als Linked Open Data) opgehaald worden.

Projectpartners

De beoogde projectpartners in dit project zijn organisaties die over één of meerdere digitale bronnen beschikken die relevant zijn voor dit project.

De projectpartners zullen gevraagd worden of het projectteam ten behoeve van de Personenportal:

- Toegang kan krijgen tot gegevens over personen, die direct betrokken waren bij de Tweede Wereldoorlog. Toegang kan op vele manieren: er kan zowel gewerkt worden met een Excel export als een verbinding met een API. Wij zullen per instelling bespreken op welke wijze de gegevens geleverd kunnen worden.
- Per persoon streven wij ernaar toegang te kunnen krijgen tot naam/namen, geboortedatum, geboorteplaats, sterfdatum, sterfplaats, uw lokale Identifiers. Als de projectpartner meer data wenst te delen dan zal dat uiteraard in de personenportal opgenomen worden!
Overigens nemen we ook minder gegevens aan als de gevraagde gegevens er niet (allemaal) zijn, maar dan wordt identificatie op het niveau van het individu lastiger.

¹ Uit de gesprekken met instellingen begrepen we dat zij zelf vaak al erg veel reacties vanuit het publiek krijgen: zij lijken geen behoefte te hebben aan mogelijkheden voor reageren in de personenportal. Het wel of niet inbouwen van een 'Reageren'-optie in de personenportal wordt nog met de deelnemende instellingen besproken.

- Ook ontvangen wij graag toegang tot of een overzicht van erfgoedcollecties waar persoonsgegevens in voorkomen. Wij denken dan aan archief-, museum of bibliotheekcollecties.
- Tijd om mee te werken en na te denken over de juridische aspecten van toegang tot persoonsgegevens.
- Tijd om gegevens aan te leveren.
- Tijd om mee te denken over de web toegang tot de personenportal en om mee te testen als de portal gebouwd gaat worden.
- Communicatie over dit project.

Uitwerking

Individen

Volgens de wet mogen er per geboortedatum, per gemeente nooit meerdere mensen met dezelfde combinatie van voor- en achternaam ingeschreven worden. Administratief zouden individuen dus altijd onderscheiden moeten kunnen worden als je beschikt over voornamen, achternaam, geboortedatum en geboorteplaats.

In de praktijk beschikken we helaas vaak vanuit bronnen over minder gegevens en zal er over 'waarschijnlijkheid' gesproken moeten worden. Een nadere uitwerking is opgenomen aan het einde van dit plan.

Identifiers

Ieder individu krijgt in de personenportal een eigen unieke Identifier. Ook zullen Identifiers of URI's (duurzame webadressen) van de bronnenleveranciers overgenomen worden: zo weten we wie in welke bron genoemd wordt (in alle naam varianten).

RDF-datamodel

Voor het vastleggen van persoonsgegevens, familierelaties en gebeurtenissen (Events) in de levens van de personen gebruiken we Linked Open Standaarden conform RDF/het Semantisch Web Model.

Zonder te veel in technische details te vervallen: RDF is een open, machine leesbare manier om informatie te beschrijven².

Ieder stukje informatie, bijvoorbeeld een beroep, wordt opgedeeld in drie onderdelen:

- Subject: de bron die beschreven wordt
- Predicaat: kenmerk van die bron
- Object: de waarde die beschreven wordt

```
<rdaGr2:professionOrOccupation xmlns:rdaGr2=http://rdvocab.info/ElementsGr2/>bronsgieter</rdaGr2:professionOrOccupation>
```

Zodra gegevens uit heel verschillende bronnen en databases gemapt zijn naar de gekozen RDF-standaarden zijn ze voor computers leesbaar en kunnen er webinterfaces op gebouwd worden, zoekvragen (Query's) op los gelaten worden enzovoorts.

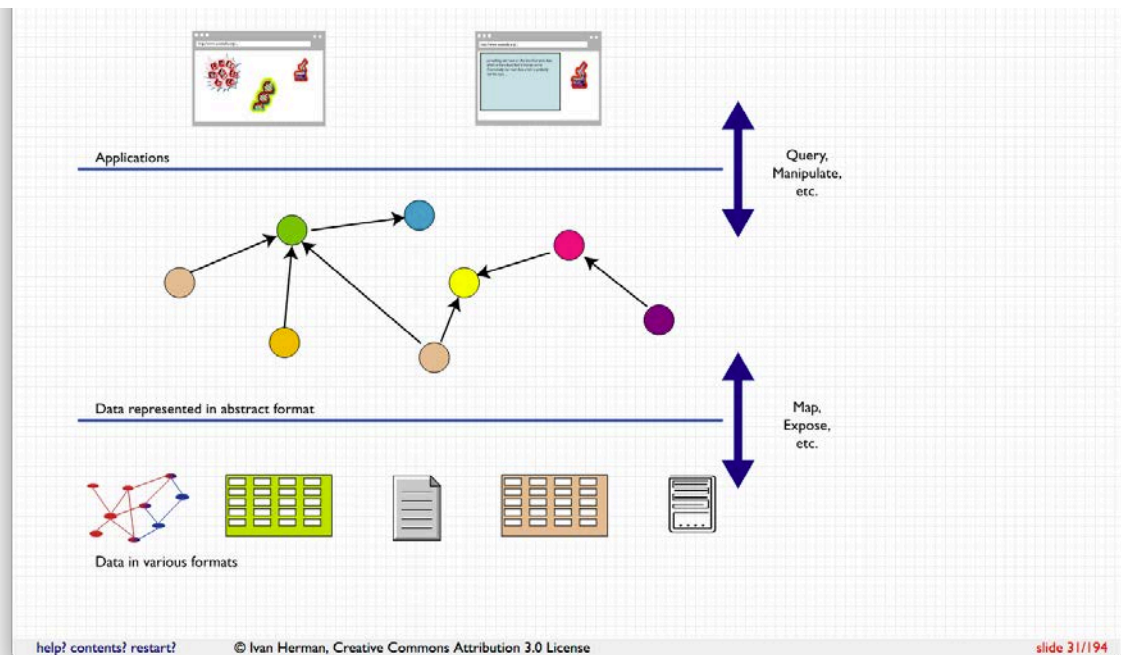
Zoals je hierboven ziet is het stukje RDF-informatie nog enigszins voor mensen leesbaar: RDF is géén software of applicatie: het is een heel gestructureerde, tekstuele, op webtechnieken gebaseerde manier om informatie voor machines toegankelijk te maken.

De standaarden (Vocabularies) die wij gekozen hebben zijn goed gedocumenteerd en online, open beschikbaar. De standaarden waar wij gebruik van willen maken zijn gangbaar binnen de erfgoedwereld (Europeana en Collectie Nederland maken er bijvoorbeeld gebruik van) en hebben bewezen dat ze door computers ingelezen kunnen worden.

Wij zullen gebruik maken van SKOS, BIO Vocab, SEM en ElementsGr2.

² Een heel duidelijke en eenvoudige uitleg over RDF is na te kijken op:
[https://www.w3.org/People/Ivan/CorePresentations/RDFTutorial/Slides.html#\(1\)](https://www.w3.org/People/Ivan/CorePresentations/RDFTutorial/Slides.html#(1))

De middenlaag (*data represented in abstract format*) in onderstaande afbeelding is de RDF laag.



Interface

Bovenop de middenlaag zal het Netwerk Oorlogsbronnen een interface bouwen die op eenvoudige doch grondige wijze toegang biedt tot de Personen. Er zal gezocht kunnen worden op persoonsnamen en naam varianten.

In samenspraak met de partners in dit project wordt vastgesteld op welke andere gegevens (plaatsen, data, beroepen, soort oorlogsdeelnemer) gezocht zal kunnen worden.

Per persoon wordt een lijst getoond met instellingen met gegevens. Wij presenteren uiteraard het liefst links naar de digitale bronnen online, maar zijn die niet beschikbaar dan komt er een melding dat er gegevens bij Instelling x beschikbaar zijn.

Soms kunnen gegevens niet openbaar beschikbaar gesteld worden, maar wel met collegae gedeeld. In dit project wordt onderzocht welke bronnen op welke wijze via een login beschikbaar gemaakt kunnen worden.

Voorbeeld van Personenportal: CDEC Shoah website. Personen worden aan Shoa Events gekoppeld en de documentatie over vervolging en bevrijding worden beschreven. Helaas is dit nog maar een kleine demonstrator.

Op de portal van Oorlogsbronnen zal per object ook een lijst met namen verschijnen van mensen die in die specifieke bron genoemd worden (of via metadata aan die bron gekoppeld zijn).

De gegevens in de personenportal zijn open beschikbaar en de deelnemende instellingen in dit project kunnen de gegevens binnenhalen in de eigen systemen of eigen interfaces op de personenportal ontwikkelen.

Communicatie

Helderheid over de stand van zaken is voor alle betrokkenen (projectteamleden en projectpartners) in dit project noodzakelijk. De projectleider en een centraal communicatiepunt vanuit NOB zorgen daarvoor. Om updates naar betrokkenen zo informatief mogelijk te maken vraagt het centrale communicatiepunt waar nodig om input van projectteamleden en /of -partners.

Natuurlijk is het gewenst als juist projectpartners aandacht geven aan het project via eigen communicatiekanalen als nieuwsbrieven, websites en sociale media. Belangrijk is daarbij wel dat de communicatie naar buiten plaatsvindt nadat betrokken partijen akkoord hebben gegeven. Dit om te voorkomen dat projectteamleden updates uit media i.p.v. intern te lezen krijgen. Het NOB neemt initiatief tot het opzetten van een communicatiewerkgroep t.b.v. afstemming.

De NOB-website bevat een projectpagina waar de voortgang van het project wordt bijgehouden voor het publiek en andere die hiernaar willen verwijzen: www.oorlogsbronnen.nl/wo2-personenportal

Communicatiemomenten

Initialisatiefase

- Gezamenlijk persbericht start project personenportal (na accordering projectplan en consent medewerking van alle partners).*
- Projectplan (ev. aangepast) publiceren op projectpagina oorlogsbronnen.nl.**

Onderzoek en handleiding relevante wet- en regelgeving

- Verslag / blogbericht in tekst en / of beeld van de workshop juridische aspecten (door NOB).*
- Handleiding rechten publiceren op projectpagina oorlogsbronnen.nl (na commitment van alle partners).**

Testsets personendata

- Ev. blogbericht(en) over de ontwikkeling van het datamodel voor proefsets en / of het opbouwen van de triple store.*
- Updatebericht op partners toegespitst n.a.v. resultaten uit testen proefsets (na rapportage en feedback per organisatie).*
- Updatebericht op projectpagina oorlogsbronnen.nl: algemene bevindingen testen proefsets (na informeren genoemde partners).

Collecties matchen

- Updatebericht op projectpagina oorlogsbronnen.nl (na informeren partners): uitgevoerde acties i.v.m. privacy en oplevering semantische omgeving.**
- Ev. blogbericht over het matchingsproces van collecties en het trainen van matchingstools (na informeren genoemde partners).*
- Updatebericht op partners toegespitst n.a.v. terug geleverde, verrijkte dataset.*
- Updatebericht op projectpagina oorlogsbronnen: algemene bevindingen matchingsproces.

Presentatie persoonsgegevens

- Testen personenportal ev. goed moment voor een vermelding via sociale media kanalen.

Testen onderzoekers

- Wervingsbericht voor externe gebruikers voor een testronde van de personenportal.*
- Verslag / blogbericht in tekst en / of beeld van de workshop voor externe gebruikers (door NOB).*

Afronding en oplevering

- Gezamenlijk persbericht over de oplevering en het gebruik van de personenportal (ev. na oplevering gebruiksrapportages partners).*

* Partners kunnen deze verspreiden via eigen kanalen.

** Partners kunnen hier ev. naar verwijzen via eigen kanalen.

Privacy wet- en regelgeving en overige juridische aspecten

Aan een personenportal, zoals beoogd te realiseren met dit project, kleven juridische bezwaren. De belangrijkste wettelijke beperkingen hebben betrekking op privacywetgeving. Als personen, opgenomen in de personenportal, nog in leven zijn en geen toestemming hebben gegeven voor publicatie van hun persoonsgegevens dan kunnen wij daarmee de Wet Bescherming Persoonsgegevens overtreden. https://autoriteitpersoonsgegevens.nl/sites/default/files/downloads/rs/rs_20071211_persoonsgegevens_op_internet_definitief.pdf

Aan de andere kant zegt ICT-jurist Arnoud Engelfriet: “als je dood bent heb je geen privacy meer. Je mag namen en informatie publiceren over deze personen zonder nadere toestemming of een belangenafweging. Alleen als de publicatie óók de privacy raakt van nog levende personen (bijvoorbeeld de kinderen of kleinkinderen), kan alsnog toestemming nodig zijn.”
Zie: <http://blog.iusmentis.com/2011/09/09/geldt-de-privacywet-ook-voor-overleden-personen/>

Voor geboorte- en huwelijksaktes gelden wettelijke termijn waarna de aktes (digitaal) publiek toegankelijk gemaakt mogen worden.

Bij grote databanken, zoals de personenportal zal worden, is het ondoenlijk om op individueel niveau toestemming voor publicatie te vragen. Wat we wel kunnen en moeten doen is de personenportal aanmelden bij het College Bescherming Persoonsgegevens. En daarnaast filters en logins in het systeem in bouwen om te voorkomen dat we gegevens over levende personen naar buiten lekken, terwijl we deze gegevens intern wel beschikbaar willen hebben. Tot slot moeten we onderzoeken hoe we de gegevens binnen de projectgroep kunnen delen zonder regels te overtreden.

Naast de wet op Bescherming Persoonsgegevens hebben we in dit project wellicht ook nog te maken met portretrecht en auteursrechten. Bij overlijden vervalt een persoonsrecht op privacy, het portretrecht nog niet: dat geldt tot 5 jaar na overlijden. En auteursrecht vervalt 70 jaar na overlijden van de auteur of opheffing van de auteursrechten beherende organisatie.

Bij aanvang van het project dient het juridisch kader helder gedocumenteerd te worden. Wij huren een adviseur in (waarschijnlijk van Kennisland) om een juridische handleiding te schrijven. Aansluitend op de handleiding wordt er ook voorlichting voor de instellingen georganiseerd.

Naast juridische kaders, maken de meeste organisaties ook keuzes ten aanzien van online publiceren op grond van moreel-ethische overwegingen. In dit project worden de keuzes en argumenten van alle instellingen meegenomen. In samenspraak met de instellingen wordt bepaald welke gegevens op welke wijze gedeeld worden.

Projectteam

Management – Stuurgroep: vertegenwoordigers partnerorganisaties personenportal

Management - Opdrachtgever: Puck Huitsing (NOB)

Management - Opdrachtnemer: Edwin Klijn (NOB)

Uitvoering - Projectleider: Lizzy Jongma (NOB)

Uitvoering - Ontwikkelen Linked Open Data backend: Inhuur (NOB)

Uitvoering - Ontwikkelen Frontend personenportal: Inhuur (NOB)

Uitvoering - Matching personen en data modellering: Inge van Stokkom (NOB)

Uitvoering – Communicatie: Tessa Free (NOB)

Uitvoering - Projectpartners:

- Leveren gegevens
- Meewerken aan onderzoek handleiding Wet- en Regelgeving
- Controleren resultaten testsets
- Controleren resultaten definitieve sets
- Testen Personenportal en leveren feedback
- Meewerken aan communicatie over dit project
- Terughalen gewenste gegevens

Doorlooptijd

De doorlooptijd van dit project is 9 maanden: februari 2017 – november 2017

Projectstappen

Initialisatie fase

Aanvullen, aanpassen projectplan.

Wie: Projectleider, projectmanagement

Accorderen projectplan.

Wie: Stuurgroep

Consent medewerking project Personenportal per organisatie.

Onderzoek en handleiding relevante wet- en regelgeving

Workshop waarin juridische aspecten, instituutsafspraken en ethische aspecten van persoonsgegevens besproken en getoetst, afgestemd worden.

Wie: allen

Handleiding voor dit project. Nadruk ligt op privacywetgeving, maar ook auteursrecht en portretrecht dienen beschreven te worden.

Wie: Adviseur Kennisland, projectleider

Afstemming afspraken met de partnerorganisaties. Commitment van alle partners.

Wie: projectleider

Testsets personendata

Vaststellen en (door-)ontwikkelen datamodel proefset.

Wie: Projectleider

Opbouwen semantische omgeving (triple store).

Wie: Backend ontwikkelaar

Inrichten omgeving voor inladen proefsets.

Wie: projectleider

Aanleveren proefset persoonsgegevens per organisatie.

Wie: Partnerorganisaties

Modelleren proefset per organisatie, schrijven handleiding.

Wie: inhuur NOB

Inrichten en *tweaken* (verbeteren) technieken voor matching.

Wie: inhuur NOB
Proefmatchen op reeds beschikbare bronnen in Oorlogsbronnen.
Wie: inhuur NOB
Rapportage en feedback per organisatie.
Wie: projectleider, inhuur NOB, partnerorganisatie

Test API/strategie voor updaten en terugleveren data

Met partnerorganisaties vaststellen wie behoefte heeft aan verrijkte data, hoe de data terug geleverd kan worden. Met partnerorganisaties vaststellen of/hoe aangevulde gegevens periodiek doorgegeven worden aan personenportal.

Wie: Projectleider, partnerorganisaties
Een of meerdere APIs worden in beta opgeleverd om verrijkte data terug te leveren.
Wie: Projectleider, partnerorganisaties

Testen beta APIs.
Wie: Inhuur NOB, partnerorganisaties

Collecties matchen

Aanmelden personenportal bij CBP en andere acties ivm privacy.
Wie: projectleider

Opleveren semantische omgeving (triple store).
Wie: Backend ontwikkelaar

Leveren datasets.
Wie: projectpartners

Matchen collecties.
Wie: Inhuur NOB

Iteratief verbeteren en trainen tools voor matching.
Wie: Inhuur NOB

Verwerkte datasets inladen in de semantische omgeving.
Wie: Inhuur NOB

Implementatie update procedures voor persoonsgegevens.
Wie: projectleider, Backend ontwikkelaars.

Terugleveren data per projectpartner en rapportage over de dataset.
Wie: Inhuur NOB, partnerorganisatie (projectleider)

Presentatie persoonsgegevens

Ontwikkelen scenario's met open en beveiligde omgevingen i.v.m. wensen en regels op gebied van privacy en bescherming data.

Wie: projectleider, projectpartners

Ontwikkelen frontend personenportal.
Wie: Frontend ontwikkelaar

Koppelen semantische omgeving aan frontend.
Wie: Frontend en backend ontwikkelaars

Inregelen en testen beveiligde gedeelte personenportal.
Wie: Frontend en backend ontwikkelaars

Ontwikkelen testscenario's.
Wie: Inhuur NOB

Testen personenportal.
Wie: Allen (testteam samengesteld vanuit projectgroep)

Iteratief verbeteren personenportal.
Wie: Frontend en backend ontwikkelaars

Inregelen Google Analytics voor rapportages en gebruik personenportal.
Wie: Ontwikkelaars en Communicatiemedewerker NOB

Testen onderzoekers, onderzoekenden

Werven externe gebruikers voor testen personenportal.

Wie: Allen

Workshop/Soft launch voor geselecteerde groep externe gebruikers.

Wie: Projectleider, Frontend en backend ontwikkelaars

Feedback analyseren.

Wie: Inhuur NOB

Verbeteren personenportal. Ontwikkelen stappenplan voor verbeteringen.

Wie: Frontend en backend ontwikkelaars, projectleider.

Go - no go besluit personenportal.

Wie: allen.

Afronding en oplevering

Oplevering rapportages.

Rapportages gebruik (Analytics) voor partners beschikbaar.

Update procedures data lopen.

Uitwerking Naamvarianten

Herman Isaïc de Vries, 14-08-1891 is waarschijnlijk Heiman Izaak de Vries, Harlingen 14-08-1891. Maar is het ook H. de Vries of H.I. de Vries (zonder verdere verwijzingen naar geboorteplaats of –jaar)? We zijn ons ervan bewust dat er in veel bronnen te weinig gegevens te vinden zijn om individuen (onomstotelijk) te kunnen identificeren.

In dit project zal gebruik gemaakt worden van OpenRefine, Spinqe Desk, Python en de programmeertaal R. Om datasets te matchen. Variaties in namen, plaatsen en data zullen door de software gedetecteerd worden. Op basis van de gedetecteerde verschillen zullen regels ontwikkeld worden. Bijvoorbeeld: Heinrich is heel waarschijnlijk Hendrik maar nooit Heidi.

Ook zullen er wegingen ingesteld worden: variaties in sterfdatum wegen wellicht minder zwaar dan variaties in geboorteplaats. En zo zullen er met behulp van wegingen waarschijnlijkheden geformuleerd worden:

Dataset A / Person X : 80% : Dataset B / Person Y [**date of death: 8-4-1943** | **date of death: 12-5-1943**].

In de personenportal komt dan:

Persoon Q010

[de gegevens waar alle bronnen eenduidig over zijn]

Dataset A / Person X

[de afwijkende gegevens]

Dataset B / Person Y

[de afwijkende gegevens]

Aan de beheerders van beide datasets zullen de verschillen uiteraard voorgelegd worden, want wellicht beschikt de ene dataset over actuelere gegevens dan de andere en kan een van de datasets gelijk getrokken worden.

Verschillende namen

Geïdentificeerde individuen zijn zowel op hun “voorkeursnaam” als op naamvarianten in bronnen te vinden (“niet-voorkeursnaam”). De voorkeursnaam van een individu is:

Doop- of Geboortenamen: Heiman Izaak

Achternaam: de Vries

Notatie: Heiman Izaak de Vries

En bij gehuwde vrouwen:

Doop- of Geboortenamen: Helena

Achternaam Echtgenoot: Bragmeijer

Achternaam: Krant

Notatie: Helena Bragmeijer-Krant

Naam varianten

Naast de voorkeursnaam (de volledige set officiële voor- en achternamen voor zover te achterhalen) worden ook naamvarianten van personen vastgelegd als het een roep- of schuilnaam betreft. Bijvoorbeeld:

Portal-nummer: 000000001

Doop- of Geboortenamen: Helena

Achternaam Echtgenoot: Bragmeijer

Achternaam: Krant

Voorkeursnaam: Helena Bragmeijer-Krant

Naamvariant: Lena Bragmeijer-Krant

Naamvariant: H. Bragmeijer-Krant

Geboorteplaats: Utrecht

Geboortedatum: 26-10-1903

Sterfplaats: Vught

Sterfdatum: 16-01-1944

Portal-nummer: 000000002
Doop- of Geboortenamen: Joseph Marie
Achternaam: van de Mortel
Voorkeursnaam: Joseph Marie van de Mortel
Naamvariant: Joseph Maria van de Mortel
Naamvariant: Joost van de Mortel
Naamvariant: J. M. van de Mortel
Geboorteplaats: Tilburg
Geboortedatum: 05-04-1919
Sterfplaats: Vught
Sterfdatum: 10-08-1944