



SBI R Oproep, 24 juli 2019

“Historische transcripties toegankelijk met AI **voor Nationaal Archief**”

Openingsdatum: 1 augustus 2019
Sluitingsdatum: 23 september 2019

Budget voor prototypes: €200.000

Ministerie van Onderwijs,
Cultuur en Wetenschap

Ministerie van Economische
Zaken en Klimaat

Nationaal Archief

Rijksdienst voor ondernemend
Nederland

SDG 16 (sterke publieke diensten)

Het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW), in het bijzonder het Nationaal Archief (NA), en het Ministerie van Economische Zaken en Klimaat (EZK) dagen ondernemers uit om nieuwe producten en diensten te ontwikkelen voor het toepassen van Artificiële intelligentie (AI).

Nederland heeft veel te winnen bij het optimaal benutten van de mogelijkheden van digitalisering en sleuteltechnologieën¹. Een van deze sleuteltechnologieën is artificiële intelligentie (AI). AI refereert aan systemen die intelligent gedrag vertonen door hun omgeving te analyseren en in zekere mate zelfstandig actie te ondernemen om specifieke doelen te bereiken².

Het Nationaal Archief heeft de laatste jaren veel gedigitaliseerd. Historische handgeschreven archieven zijn als plaatjes beschikbaar op internet maar de informatie in deze documenten is niet doorzoekbaar. Daar willen we een eind aan maken.

SBIR³ is een innovatie-competitie om ondernemers uit te dagen om nieuwe producten of diensten te ontwikkelen gericht op de aanpak van een maatschappelijk vraagstuk. Wij zijn op zoek naar concreet bruikbare toepassingen, die onder andere kunnen worden gebruikt door (semi-) publieke organisaties. De overheid is daarbij een potentiële (maar geen exclusieve) inkoper.

¹ <https://www.topsectoren.nl/missiegedreven-innoveren>

² Kamerstukken II, 2017–2018, 22 112, nr. 2578, p. 2 (brief van de minister van Buitenlandse Zaken van 1 juni 2018: Fiche: Mededeling Kunstmatige intelligentie voor Europa).

³ De afkorting SBIR komt van het Amerikaanse Small Business Innovation Research program. Het Amerikaanse SBIR programma is alleen toegankelijk voor kleine bedrijven. Het Nederlandse SBIR programma is toegesneden op kleinere ondernemingen, maar staat open voor iedereen, ook voor grote ondernemingen.

1. Doel van deze SBIR-Innovatiecompetitie

In het innovatieve project 'De ijsberg zichtbaar maken' (bijlage 1), transcriberen we met behulp van machine learning twee miljoen scans van historische handgeschreven documenten. We gebruiken hiervoor het transcriptieplatform [Transkribus](#).

We selecteerden 1 miljoen scans van het archief ⁴van de Verenigd Oost-Indisch Company (VOC) uit de 17e en 18e eeuw van het Nationaal Archief. Van de Regionaal Historische Centra (RHC) selecteerden we 1 miljoen scans van notariële archieven uit de 19e eeuw.

Doel van deze SBIR-Innovatiecompetitie is het vinden van oplossingen waardoor deze grote hoeveelheid transcripties voor een onderzoeker hanteerbaar wordt.

Belangrijk is daarbij dat de onderzoeker vindt wat hij zoekt. Vaak zoekt hij naar een persoonsnaam, een naam van een locatie of naar een gebeurtenis. Hoe kan hij zo goed mogelijk worden geholpen bij zijn zoektocht? Hoe kan de presentatie of het interactiemodel daarbij helpen?

Bij het *full-text* doorzoeken van de transcriptiedata krijgt de onderzoeker een overvloed aan resultaten te verwerken. Hoe ga je met onder andere AI de onderzoeker helpen door de data slim te bewerken, te indexeren, te ordenen en te verbinden?

Van elk archief is er een beschrijving beschikbaar. We noemen zo'n beschrijving een toegang. Deze beschrijving is op de website niet automatisch gekoppeld aan de transcripties maar is wel nodig om op een gestructureerde manier de zoekresultaten te kunnen filteren. Zo staat er in wie het archiefstuk heeft gemaakt, wat voor type archiefstuk het is en wanneer en waar het is gevormd. Hoe ga je deze informatie gebruiken om de onderzoeker te helpen?

Welke oplossing heb je voor transcriptiefouten, spellingsvarianten en oud Nederlands?

In de 'Toelichting op het doel van deze SBIR-Innovatiecompetitie' (bijlage 2) staat aanvullende informatie over het archief dat wordt getranscribeerd, de beschikbare datasets en de hierboven genoemde uitdagingen.

Let op:
De eindoplossing (niet persé het prototype) moet voldoen aan de Solution Architecture ([bijlage 3](#)).

⁴ Een archief is een geheel aan archiefstukken ontvangen en/of gemaakt door een persoon of organisatie

2. Procedure

SBIR is een open competitie voor iedere marktpartij die innovatieve (technologische) oplossingen voor maatschappelijke vraagstukken kan ontwikkelen. De SBIR-systematiek kent twee fasen:

1. Fase 1: haalbaarheidsonderzoek (oplossingen exploratie).
2. Fase 2: toegepast onderzoek en ontwikkeling (prototyping).

Deze oproep kent geen fase 1. Het Nationaal Archief heeft onlangs uitgebreid voorbereidend onderzoek gedaan:

- De ijsberg zichtbaar maken (bijlage 1): het projectplan van het Nationaal Archief voor het toegankelijk maken van historische handschriften met automatische handschriftherkenning.
- Toelichting op het doel van deze SBIR-Innovatiecompetitie (bijlage 2): aanvullende informatie over het archief dat wordt getranscribeerd, de beschikbare datasets en de belangrijkste uitdagingen voor het zoeken en tonen van transcripties.
- Solution architectuur (bijlage 3): de beschrijving van het architectuurkader waarbinnen de functionaliteit voor het zoeken en tonen van transcripties wordt gerealiseerd. Het geeft samen met de functionele en de niet-functionele specificaties richting aan de uiteindelijk te realiseren Software Architectuur.
- **Persona's nieuwe website** - versie SBIR oproep (bijlage 4): **Persona's zijn in het leven geroepen om te ondersteunen bij de ontwikkeling van de nieuwe website, zowel conceptueel als bij de realisatie. De persona's vertegenwoordigen klantgroepen van het Nationaal Archief. In deze bijlage zijn enkel de persona's opgenomen die relevant zijn voor de SBIR oproep.**

Een onafhankelijke commissie zal adviseren welke voorstellen voor producten en diensten (fase 2) het beste aan de criteria voldoen. De beste voorstellen krijgen een opdracht om een prototype te ontwikkelen.

Het Nationaal Archief is een agentschap van het ministerie van OCW. Binnen het ministerie van OCW valt het Nationaal Archief onder de verantwoordelijkheid van de Directeur-Generaal Cultuur en Media.

De Rijksdienst voor Ondernemend Nederland (RVO) voert namens het Nationaal Archief deze SBIR uit. RVO is een agentschap het ministerie van EZK.

3. Budget

Er is in **totaal een budget van maximaal € 200.000** (inclusief BTW) beschikbaar voor het ontwikkelen van drie (3) prototypes.

Het maximum per prototype **bedraagt € 66.666,-** (inclusief BTW).

Na afronden van het SBIR traject beoogt het Nationaal Archief om met één (1) partij een vervolgtraject in te zetten voor het verder doorontwikkelen van het prototype. **Hiervoor is een budget van €385.000 beschikbaar.**

De doorontwikkeling vindt plaats met een partij die een prototype heeft ontwikkeld in het kader van het SBIR traject. Vanuit het oogpunt van zorgvuldigheid en transparantie zal het NA dit voornemen, voorafgaand aan opdrachtverstrekking, kenbaar maken aan marktpartijen.

4. Beoordeling

De beoordeling vindt plaats conform de in de SBIR handleiding voor ondernemers beschreven procedure: <https://mijn.rvo.nl/sbir-innovatie-in-opdracht>

En aan de hand van onderstaande beoordelingscriteria:

Beoordelingscriteria

Bij de beoordeling is in totaal maximaal 100 punten te behalen. Per criterium is maximaal het volgende aantal punten toe te kennen:

1. Impact: 50
2. Technische haalbaarheid: 40
3. Economisch perspectief: 10

Impact

Een voorstel scoort hoger op impact als deze beter aansluit bij de doelstelling van de oproep.

Voor deze oproep zijn voor het criterium 'Impact' de volgende aspecten van belang. Een offerte scoort hoger naarmate:

1. de onderzoeker met behulp van het product het 'oerwoud' van transcriptiedata doelmatiger en effectiever kan doorzoeken;
2. het systeem rekening houdt met specifieke uitdagingen zoals transcriptiefouten, alternatieve spellingen en oud Nederlands;
3. er door nieuwe ordening, sortering of alternatieve presentatie van resultaten verbanden intuïtiever of levendiger kunnen worden overgebracht en
4. meer aannemelijk is dat de oplossing technisch robuust is (veilig, betrouwbaar, responsief).

Technologische haalbaarheid

Bij het criterium '(Technologische) haalbaarheid' zijn de volgende aspecten van belang:

1. Kwaliteit van de offerte (plan van aanpak inclusief afstemming met klant en de onderbouwing door een visie.)
2. De mate van technologische vernieuwing.
3. De duurzaamheid van de oplossing; de mate waarop de oplossing kan meebewegen met nieuwe ontwikkelingen. Dit kan onderbouwd worden door toe te lichten hoe open standaarden en een modulaire architectuur worden toegepast als voorbereiding op aansluiting van toekomstige ontwikkelingen.
4. De mate van
 - a) overdraagbaarheid van beheer en doorontwikkeling aan het Nationaal Archief,
 - b) integratie in het bestaande technische landschap van het Nationaal Archief en
 - c) de bereidheid om de ontwikkelde software onder opensource licentie te plaatsen.

Economisch perspectief

Bij het criterium 'Economisch perspectief' wordt gekeken naar:

1. Ondernemerschap;
2. Het (opensource-) verdienmodel;
3. Verwachte betrokkenheid en interesse van (potentiële) afnemers van het concept;
4. De voorgenomen aanpak en strategie om het concept een plaats in de markt te bezorgen;
5. De kansen in de Nederlandse markt (herhaalpotentieel) en
6. Kansen voor export.

5. Informatiebijeenkomst

Op dinsdag 3 september 2019 zal bij Nationaal Archief in Den Haag een informatiebijeenkomst worden georganiseerd.

Het programma van de bijeenkomst ziet er globaal als volgt uit:

- Welkom – Nationaal Archief
- Uitleg over de Procedure - RVO
- Uitleg over inhoud van deze SBIR - Nationaal Archief
- Solution architectuur- Nationaal Archief
- User Journey - Nationaal Archief
- Afsluiting – RVO
- Mogelijkheid tot netwerken

U kunt zich hiervoor uiterlijk 28 augustus aanmelden via de mail sbir@rvo.nl met vermelding van uw gegevens (telefoonnummer en organisatie). U dient bij het bezoek aan het Nationaal Archief een geldig legitimatiebewijs mee te nemen.

6. Informatie en contact

Alle informatie over deze SBIR-competitie en relevante SBIR-documenten vindt u op de volgende websites (voor openen van de snelkoppelingen ctrl-toets ingedrukt houden en met linkermuisknop aanklikken):

<https://www.rvo.nl/subsidies-regelingen/sbir> en <http://www.tenderned.nl>

Heeft u vragen met betrekking tot deze SBIR-competitie dan kunt u deze (tot maximaal 10 dagen voor sluitingsdatum) sturen naar: sbir@rvo.nl.

6a. Technische Informatie

Via de volgende links komt u bij de datapakketten die behulpzaam zullen zijn bij uw onderzoek en om een prototype te ontwikkelen.

<http://sbir.datapakket.nl/>

7. Indienen offertes (projectvoorstellen)

U dient via e-mail uw offerte in te dienen. De offerte dient uiterlijk op dinsdag 23 september 2019 om 13.00 uur in het bezit te zijn van RVO.nl via sbir@rvo.nl.

Let op: te laat ingediende offertes worden niet meegenomen in de beoordeling. De mailboxserver kan een omvang van 10MB of meer niet aan (meerdere mails of gebruik maken van een dienst als We-transfer biedt dan een oplossing).

Een volledige SBIR-offerte fase 2 bestaat uit:

- Het ingevulde en ondertekende SBIR-formulier
- Het projectplan (inclusief begroting en openbare samenvatting)
- De managementsamenvatting
- Het businessplan

Daarnaast wordt het meezenden van een Video-pitch (vlog) van maximaal 3 minuten in MP4-formaat zeer gewaardeerd.

Alle bestanden moeten daarbij als aparte PDF-, Word- of Excelbestanden worden aangeleverd. De benodigde formats voor uw offerte zijn te vinden op:

<https://mijn.rvo.nl/sbir-innovatie-in-opdracht>

Wij raden u aan om een aantal werkdagen voor de deadline uw offerte in te dienen.

8. Planning

Openstelling tender	Donderdag 1 augustus 2019
Informatiebijeenkomst	Dinsdag 3 september 2019
Sluiting indienen fase 2 offertes	Maandag 23 september 2019, 13.00 uur
Commissievergadering (toelichten offerte door indiener)	Maandag 7 oktober 2019
Bekendmaking uitslag	Medio oktober 2019
Opdrachtverstrekking	15 oktober 2019
Deadline eindrapport en prototype	15 februari 2020

RVO.nl behoudt zich het recht voor om bijgevoegd tijdsplan indien nodig aan te passen. Dit zal tijdig aan (potentiële) opdrachtnemers worden gecommuniceerd.

- Bijlage 1: Projectplan De ijsberg zichtbaar maken
- Bijlage 2: Toelichting op het doel van deze SBIR-Innovatiecompetitie
- Bijlage 3: Solution Architectuur
- Bijlage 4: **Persona's nieuwe website** - versie SBIR oproep