



Digitale erfgoedverhalen bewaard voor de toekomst

Een onderzoek naar het beheren en
beschikbaar stellen van digitale verhalen

April 2024





Inhoud

1. Inleiding	3
Aanleiding	3
Onderzoeksvragen en doel	3
Leeswijzer	4
2. Wat zijn verhalen?	5
Kenmerken van verhalen	5
Belang van verhalen	6
Soorten en vormen van verhalen	6
3. Wat is de structuur van verhalen?	8
Informatiemodel	8
Schematische weergave	11
Voorbeeld van een verhaal volgens het informatiemodel	12
4. Kan software helpen om termen toe te kennen aan verhalen?	14
Verbindingen met termen	14
Software	14
5. Hoe verder?	17
Colofon	18

1. Inleiding

Dit hoofdstuk geeft een introductie op ons onderzoek naar erfgoedverhalen.

Aanleiding

Steeds meer erfgoedinstellingen, van klein tot groot, maken hun erfgoed digitaal toegankelijk via websites en apps. Dat is een fantastische ontwikkeling. Leerlingen, journalisten, historici en tal van andere typen gebruikers kunnen erfgoed hierdoor vinden en gebruiken.

In onze praktijk zien wij dat erfgoedinstellingen vaak twee soorten informatie aanbieden. In de eerste plaats is dat informatie over objecten in collecties, zoals metadata en afbeeldingen. In de tweede plaats zijn dat verhalen, veelal geschreven vertellingen over gebeurtenissen uit het verleden.

Met deze verhalen is iets bijzonders aan de hand. Verhalen geven kleur en betekenis aan erfgoed. Ze zijn essentieel om erfgoed te duiden, te verbinden en te verrijken. Verhalen *zijn* volgens ons dan ook erfgoed – volwaardige, opzichzelfstaande informatieobjecten¹ die, net als objecten in collecties, met zorg bewaard en ontsloten moeten worden.

Toch zien wij in onze praktijk ook dat verhalen vaak over het hoofd worden gezien. Zo is er doorgaans geen informatiebeleid voor. Erfgoedinstellingen zijn zich onbewust hoe verhalen duurzaam toegankelijk of bewaard moeten worden op de langere termijn. En soms worden verhalen beheerd in systemen die daarvoor niet zijn toegerust. Dat komt waarschijnlijk omdat er nog veel onduidelijk is over verhalen.

Bij Erfgoed Brabant² en Huis voor de Kunsten Limburg³ willen wij dat veranderen. Wij willen de verhalen die in onze provincies gemaakt worden op een doordachte, toekomstvaste manier beheren en beschikbaar stellen. Het Netwerk Digitaal Erfgoed (NDE)⁴, het landelijke samenwerkingsverband van erfgoedinstellingen, signaleert een soortgelijke behoefte bij anderen. In opdracht van het NDE hebben wij een kort onderzoek gedaan naar enkele vragen die wij hebben over verhalen. De resultaten hiervan zijn in dit paper opgenomen.

Onderzoeksvragen en doel

Bij de aanvang van ons onderzoek hebben wij drie onderzoeksvragen geformuleerd:

1. Wat zijn verhalen?
2. Wat is de structuur van verhalen?
3. Kan software helpen om termen toe te kennen aan verhalen?

Het doel van deze vragen is om meer grip te krijgen op verhalen. Dit stelt ons in staat om te bepalen welke stappen wij moeten zetten om verhalen op een duurzame manier te kunnen opslaan en toegankelijk te maken.

¹ <https://www.nationaalarchief.nl/archiveren/kennisbank/informatieobject>

² <https://www.erfgoedbrabant.nl>

³ <https://hklimburg.nl>

⁴ <https://netwerkdigitaalerfgoed.nl>

Het onderzoek is specifiek voor onze provinciale situaties en belicht onderwerpen die wij nu belangrijk vinden. Het neemt onze ervaringen met het ontsluiten van verhalen via platformen als BrabantsErfgoed.nl als vertrekpunt en houdt rekening met ons denken over het toepassen van principes van het NDE rond het houdbaar, bruikbaar en zichtbaar maken van erfgoed. Toch denken wij dat het onderzoek eveneens waardevol kan zijn voor medewerkers van andere erfgoedinstellingen. Dit paper is dan ook bedoeld voor informatiespecialisten en ict'ers van instellingen die, net als wij, aan de slag willen gaan met verhalen en daar vragen over hebben.

Leeswijzer

Het paper bestaat uit vijf hoofdstukken. Hoofdstuk 1 – dit deel – geeft een introductie op het paper. Hoofdstuk 2 bespreekt de eerste onderzoeksvraag: wat zijn verhalen? Hoofdstuk 3 gaat in op de tweede onderzoeksvraag: wat is de structuur van verhalen? Hoofdstuk 4 beantwoordt de derde onderzoeksvraag: kan software helpen om termen toe te kennen aan verhalen en ze beter vindbaar te maken? Hoofdstuk 5, tot besluit, benoemt hoe wij verder willen gaan met de resultaten van het onderzoek.

2. Wat zijn verhalen?

Dit paper gaat over verhalen. Maar wat zijn verhalen nu precies? Daarover gaat de eerste onderzoeksvraag. Hoe scherper dat antwoord wordt, hoe meer zicht we krijgen op wat dit betekent voor het beheren en publiceren ervan en wat de gevolgen zijn voor bijvoorbeeld beleid, functionaliteit van applicaties en technologie van systemen.

Om de onderzoeksvraag te kunnen beantwoorden, hebben wij verschillende medewerkers van erfgoedinstellingen in Limburg en Noord-Brabant geraadpleegd. Deze medewerkers hebben uiteenlopende praktijkervaringen met het maken en presenteren van verhalen via hun eigen digitale kanalen, zoals websites. Wij hebben ze gevraagd wat de kenmerken van verhalen zijn, wat het belang van verhalen is en welke soorten en vormen van verhalen zij onderscheiden. Hun antwoorden worden hierna per onderdeel puntsgewijs samengevat.

Kenmerken van verhalen

Het doel van dit onderdeel is om de reikwijdte van een verhaal te definiëren: wat is het en wat is het niet? In onze gesprekken met medewerkers is het lastig gebleken om een definitie van een 'verhaal' te maken waar iedereen zich in kan vinden; deze wordt abstract en onwerkbaar. Erfgoedinstellingen moeten ruimte houden om zelf te bepalen welke definitie voor hun situatie relevant is. De volgende kenmerken, die tijdens de gesprekken met de instellingen naar voren zijn gebracht, helpen om een eigen definitie te formuleren.

1. Een verhaal gaat over het verleden, niet over de toekomst. 'Verleden' is heel ruim – het kan gaan over iets dat lang geleden gebeurd is, maar ook onlangs.
2. Een verhaal is een vertelling over een gebeurtenis. Een 'gebeurtenis' kan klein of groot zijn. Bijvoorbeeld: een specifiek voorval op een bepaalde dag of 'veldslagen in Limburg door de eeuwen heen'. Daarbij maakt een verhaal de context van de gebeurtenis duidelijk. Bijvoorbeeld: in welke omstandigheden vond het plaats en hoe is het verbonden met andere gebeurtenissen?
3. Een verhaal kan over allerlei onderwerpen gaan. Het kan bijvoorbeeld een bepaalde persoon betreffen of een bepaald object, zoals een zwaard of een transistorradio.
4. Een verhaal is bij voorkeur non-fictie, maar hoeft niet waargebeurd te zijn. Het kan bijvoorbeeld een vastlegging van een lokale legende zijn. Zo'n verhaal heeft waarde omdat het iets zegt over mensen en hun kijk op de wereld.
5. Een verhaal biedt de context die nodig is om een object in een collectie te begrijpen, waarde te geven en te laten beleven. Het verhaal is misschien wel belangrijker dan het object.
6. Een verhaal is een gecureerde, samenhangende, toegankelijke vertelling voor een bepaalde doelgroep. De wijze van vertellen boeit blijvend; het spreekt mensen nu aan, maar hopelijk ook later.
7. Een verhaal heeft een digitale vorm die op het web gelezen, beluisterd of bekeken kan worden. Dit betekent dat verhalen die enkel mondeling bestaan of analogoog zijn, zoals in een tijdschrift van een heemkundekring, eerst gedigitaliseerd moeten worden.

8. Een verhaal is pas echt een verhaal als het digitaal beschikbaar blijft. De waarde van een verhaal, ook voor toekomstige lezers, kan alleen worden geborgd als de duurzame toegankelijkheid geregeld is. Dit kenmerk ten spijt, medewerkers hebben aangegeven op dit moment nog niets of weinig te doen met deze duurzaamheid.

Belang van verhalen

Het doel van dit onderdeel is om inzichtelijk te maken waarom we verhalen maken: wat is het belang ervan? De medewerkers hebben twee belangen benoemd.

1. Het eerste belang is om een groep mensen bewust te maken van hun erfgoed en zich daarover te laten verwonderen. Een 'groep' kan bijvoorbeeld bestaan uit mensen die woonachtig zijn in dezelfde provincie ('Limburgers') of regio ('Bosschenaren') of uit mensen die geïnteresseerd zijn in hetzelfde onderwerp (zoals 'Tweede Wereldoorlog'). Het achterliggende doel hiervan is om het welzijn van mensen te vergroten. Ze worden er gelukkiger van, ervaren meer sociale binding en delen een gemeenschappelijke identiteit.
2. Het tweede belang is om mensen beter te kunnen bereiken dan met objecten in collecties. Bijvoorbeeld: archiefdocumenten bevatten een schat aan informatie, maar zijn vaak alleen bruikbaar voor specialisten. Door verhalen over documenten te vertellen, kan een breder publiek aangesproken worden.

Soorten en vormen van verhalen

Het doel van dit onderdeel is om duidelijk te krijgen wat de diversiteit van verhalen is.

1. De medewerkers hebben zeven *soorten* verhalen benoemd:
 - a. Artikel, historisch verhaal of lemma: een geschiedkundig accurate weergave van een gebeurtenis. Een uitwerking hiervan is een achtergrondartikel, een weergave van een gebeurtenis met een duiding of analyse.
 - b. Biografie of levensverhaal: een levensbeschrijving van een persoon.
 - c. Interview: een vastlegging van een vraagsprek met een persoon.
 - d. Getuigenis: een herinnering van een persoon aan een gebeurtenis.
 - e. Canon: een overzicht van belangrijke gebeurtenissen uit een bepaald gebied.
 - f. Volksverhaal: een vastlegging van een legende of sage.
 - g. Historische fictie: een (vaak meeslepende) voorstelling van een gebeurtenis die gedeeltelijk op de werkelijkheid berust.
2. De medewerkers hebben drie *vormen* van verhalen benoemd:
 - a. Een 'geschreven verhaal' is een digitale of gedigitaliseerde tekst. Dit is op dit moment de belangrijkste vorm.
 - b. Een 'gesproken verhaal' is digitaal of gedigitaliseerd geluid.
 - c. Een 'verbeeld verhaal' is een digitale of gedigitaliseerde film.
 - d. Combinaties zijn er ook, zoals een tekst waarin geluid of film is opgenomen ('ge-embed'). Daarnaast is er een meer fluïde vorm: de 'digitale interactieve ervaring' die bestaat uit tekst, beeld en geluid. Deze vorm komt evenwel weinig voor bij de erfgoedinstellingen die zijn geraadpleegd tijdens dit onderzoek⁵.

⁵ Voorbeelden zijn te vinden op <https://ons.limburgsmuseum.nl>

3. Een verhaal is van een bepaalde soort en van een bepaalde vorm, zoals een geschreven biografie of een gesproken getuigenis. Het kan ook voorkomen dat een verhaal meerdere vormen heeft, zoals een biografie die zowel geschreven als gesproken is. Elke vorm kan weer subvormen hebben, zoals een geschreven biografie die als tekst en als EPUB beschikbaar is. Dit leidt tot 'meerlagigheid': de vormen zijn verschillende uitingen van hetzelfde (intellectuele/artistieke) verhaal. Hoewel deze meerlagigheid het mogelijk maakt om verhalen te groeperen en beter te presenteren, zorgt het ook voor complexiteit. De medewerkers hebben aangegeven op dit moment nog niets met deze meerlagigheid te doen en daar vooralsnog ook geen behoefte aan te hebben.

3. Wat is de structuur van verhalen?

Het antwoord op de eerste onderzoeksvraag, in het voorgaande hoofdstuk, heeft ons een beter beeld gegeven van wat verhalen zijn. Dit beeld moeten we evenwel concreter maken om verhalen digitaal te kunnen opslaan en toegankelijk te maken.

Een manier om dat te doen is door de structuur van verhalen te beschrijven. Daarover gaat de tweede onderzoeksvraag. De structuur geeft aan welke gegevens – of metadata – bij verhalen horen. Wij hebben medewerkers gevraagd welke gegevens zij idealiter willen vastleggen. Dit hebben wij vervolgens uitgewerkt in een informatiemodel⁶.

Informatiemodel

Een informatiemodel geeft aan welke informatietypen van belang zijn voor verhalen, welke attributen deze typen hebben en hoe ze onderling verbonden zijn. Een informatiemodel is conceptueel – het helpt ons om het gesprek aan te gaan over de structuur van verhalen, maar doet geen uitspraken over de precieze vorm die verhalen krijgen, bijvoorbeeld in applicaties of databases.

In ons informatiemodel onderscheiden wij tien informatietypen. Deze worden hierna kort toegelicht. Op de daarop volgende bladzijden wordt het model schematisch weergegeven en een voorbeeld gegeven van een verhaal dat volgens dit model gestructureerd is. Het model is een eerste versie.

Erfgoedverhaal

Dit informatietype beschrijft de kern van een digitaal verhaal. Het bevat de attributen waarmee een verhaal geïdentificeerd en geclassificeerd kan worden. Wij hebben het informatietype 'Erfgoedverhaal' genoemd – en niet 'Verhaal' – omdat deze naam duidelijk maakt dat het gaat om een verhaal over of gerelateerd is aan het verleden. Het informatietype is algemeen: het zegt nog niets over de vorm van het verhaal. Om ook de vorm aan te duiden, onderscheiden we drie onderliggende types: 'Geschreven erfgoedverhaal', 'Gesproken erfgoedverhaal' en 'Verbeeld erfgoedverhaal' (zie hierna).

Geschreven erfgoedverhaal

Dit informatietype beschrijft een specifieke vorm van een verhaal: geschreven. Het bevat attributen die horen bij deze vorm, bijvoorbeeld de inleiding en de inhoud. Dit informatietype kan verbonden zijn met een of meer mediaobjecten die het verhaal begeleiden, zoals afbeeldingen.

⁶ <https://www.nationaalarchief.nl/archiveren/kennisbank/informatiemodel>

Gesproken erfgoedverhaal

Dit informatietype beschrijft een specifieke vorm van een verhaal: gesproken. Het bevat attributen die horen bij deze vorm, bijvoorbeeld een beschrijving die duidelijk maakt waarover het verhaal gaat. Dit informatietype is verbonden met een of meer specifieke mediaobjecten: geluidsopnamen die beluisterd kunnen worden.

Verbeeld erfgoedverhaal

Dit informatietype beschrijft een specifieke vorm van een verhaal: verbeeld. Het bevat attributen die horen bij deze vorm, bijvoorbeeld een beschrijving die duidelijk maakt waarover het verhaal gaat. Dit informatietype is verbonden met een of meer specifieke mediaobjecten: video-opnamen die bekeken kunnen worden.

Mediaobject

Dit informatietype beschrijft een digitaal mediabestand, zoals een afbeeldings-, audio- of videobestand. Een mediaobject verwijst naar de vervaardiger van het bestand en naar de licentie waarmee het bestand toegankelijk is gemaakt.

Reactie

Dit informatietype beschrijft een tekstuele reactie die is geplaatst bij een verhaal. Reacties maken het verhaal rijker – ze kunnen bijvoorbeeld aanvullingen of correcties bevatten of een indruk geven van de mening van gebruikers. Een reactie verwijst naar de vervaardiger en naar de licentie waarmee het toegankelijk is gemaakt.

Licentie

Dit informatietype beschrijft een licentie die duidelijk maakt onder welke voorwaarden een bepaald informatietype, zoals een erfgoedverhaal, mediaobject of reactie, toegankelijk is gemaakt en hergebruikt mag worden.

Persoon of organisatie

Dit informatietype beschrijft een persoon of organisatie. Dat kan bijvoorbeeld de persoon betreffen die een verhaal gemaakt heeft of de organisatie die verantwoordelijk is voor het beheren of uitgeven ervan. Het type kan ook gebruikt worden om aan te geven over wie een verhaal gaat, bijvoorbeeld bij een biografie.

Gebeurtenis, locatie, object, onderwerp, periode

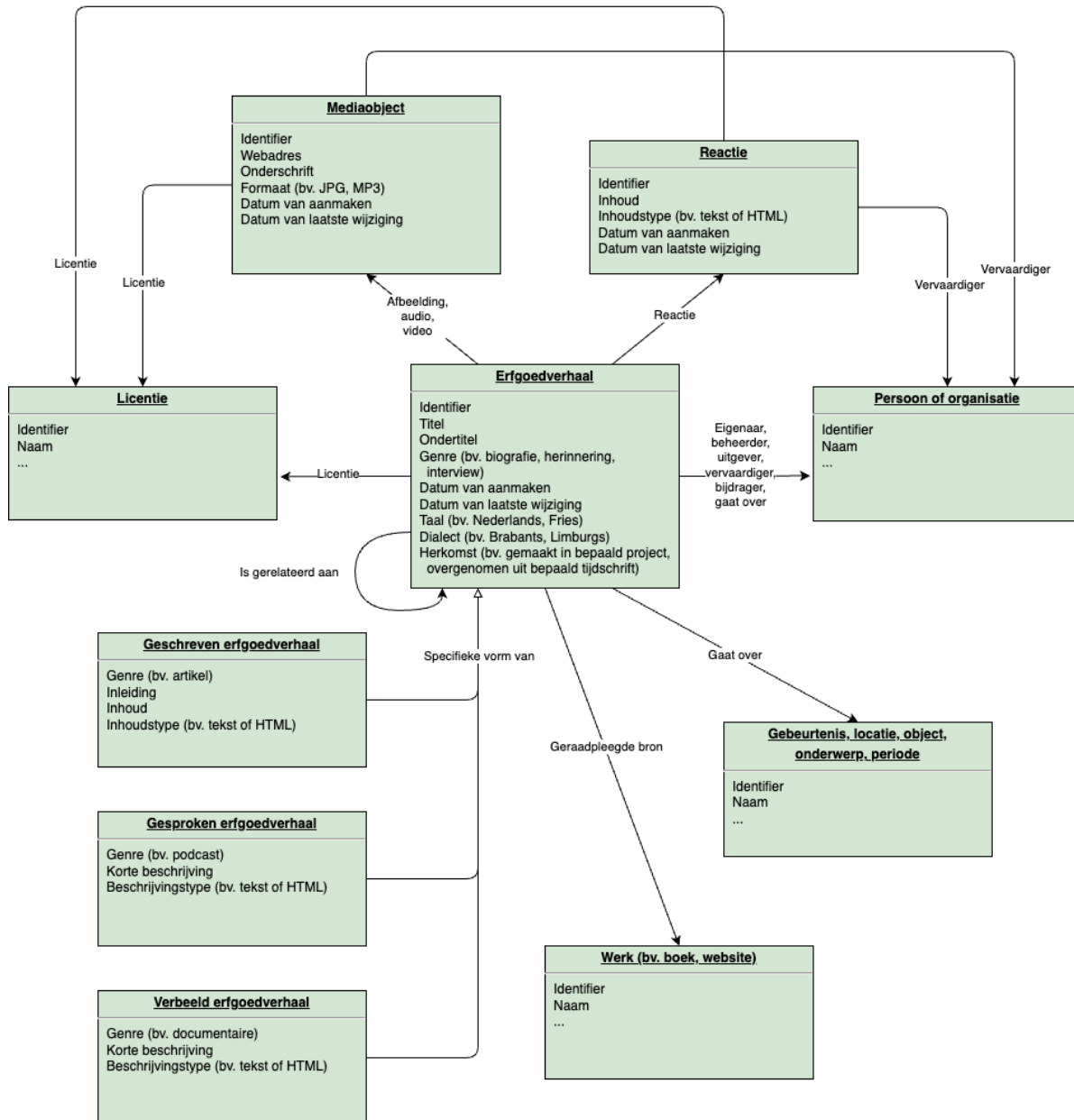
Dit is een informatietype waarin we een aantal opzichzelfstaande informatietypen hebben gegroepeerd omwille van de leesbaarheid. Dit type is cruciaal om de belangrijkste 'wie-wat-waar-wanneer'-aspecten van een verhaal te kunnen benoemen. Bijvoorbeeld om een verhaal te kunnen verbinden met een object in een collectie of om duidelijk te maken dat een verhaal over een bepaalde locatie of onderwerp gaat.

Werk

Dit is een informatietype waarin we een aantal opzichzelfstaande informatietypen hebben gegroepeerd omwille van de leesbaarheid. Dit informatietype beschrijft een werk, zoals een boek of een website. Een verhaal kan naar een werk verwijzen om duidelijk te maken dat het bij het maken van het verhaal geraadpleegd is als bron.

Schematische weergave

Een schematische weergave van het informatiemodel, met daarin de informatietypen, hun attributen en de onderlinge verbanden, wordt hierna getoond.



Figuur 1: schematische weergave van het informatiemodel

Voorbeeld van een verhaal volgens het informatiemodel

Een informatiemodel kan abstract zijn. Om het model duidelijker te maken, hebben wij een bestaand verhaal volgens het model gestructureerd. Het verhaal heet 'Vincent van Gogh op kostschool' en is gepubliceerd op de website van het Brabants Historisch Informatie Centrum (BHIC)⁷. Bij bepaalde gegevens staat '(Niet bekend)' vermeld; deze konden niet achterhaald worden op de website van het BHIC.

Informatietype: geschreven erfgoedverhaal

Identifier	(Niet bekend)
Titel	Vincent van Gogh op kostschool
Ondertitel	(Niet van toepassing)
Genre	Artikel
Inleiding	Van 1 oktober 1864 tot 31 augustus 1866 verbleef Vincent van Gogh op de kostschool van Jan Provily in Zevenbergen. Vincent zelf vertelt er in drie brieven kort iets over. Hieruit blijkt dat hij zich hier niet zo gelukkig heeft gevoeld.
Inhoud	Domineeszoon Vincent van Gogh werd op 30 maart 1853 in het Brabantse grensdorp Zundert geboren. Tot zijn 8ste jaar kreeg hij les van zijn vader (...)
Inhoudstype	HTML
Datum van aanmaken	25 april 2019
Datum van laatste wijziging	23 april 2020
Taal	Nederlands
Dialect	(Niet van toepassing)
Herkomst	Verkorte versie van https://www.bhic.nl/media/document/file/vvg-in-zevenbergen-definitief.pdf
Vervaardiger	Cor Kerstens
Eigenaar	Cor Kerstens
Beheerder	Brabants Historisch Informatie Centrum
Uitgever	Brabants Historisch Informatie Centrum
Gaat over persoon	Vincent van Gogh
Gaat over locatie	Zevenbergen
Gaat over gebouw	Voormalige kostschool, Stationsstraat 16, Zevenbergen

⁷ <https://www.bhic.nl/ontdekken/verhalen/vincent-van-gogh-op-kostschool>

Gaat over periode	1864-1866
Gaat over onderwerp	kostschool, onderwijs
Licentie	(Niet bekend)
Geraadpleegde bron	(Niet bekend)

Informatietype: mediaobject

Identificer	https://www.bhic.nl/memorix/images/search/detail/31c4cc62-45fb-11e3-8624-e7bda1721fcf/media/26097b4e-deca-d99c-4c2d-2598d67b61a0
Webadres	https://webservices.picturae.com/mediabank/media/31c4cc62-45fb-11e3-8624-e7bda1721fcf/downloadoriginal/26097b4e-deca-d99c-4c2d-2598d67b61a0?apiKey=24c66d08-da4a-4d60-917f-5942681dcaa1
Onderschrift	De voormalige kostschool, Stationsstraat 16 in Zevenbergen (BHIC, coll. provincie Noord-Brabant, fotonr. PNB001076928)
Formaat	JPEG
Datum van aanmaken	1 oktober 1989
Datum van laatste wijziging	(Niet bekend)
Vervaardiger	Leeuwen, Wies van (Provincie Noord Brabant)
Licentie	(Niet bekend)

Informatietype: reactie

Identificer	(Niet bekend)
Inhoud	Wat een mooi verhaal, Cor! Ontroerend en aangrijpend om te lezen hoe (...)
Inhoudstype	Tekst
Datum van aanmaken	1 mei 2019 om 14:46
Datum van laatste wijziging	1 mei 2019 om 14:46
Vervaardiger	Marilou Nillesen
Licentie	(Niet bekend)

4. Kan software helpen om termen toe te kennen aan verhalen?

De derde en laatste onderzoeksvraag gaat over het toekennen van termen aan verhalen om ze beter vindbaar te maken.

Verbindingen met termen

Tijdens gesprekken met medewerkers van erfgoedinstellingen is benoemd dat het van groot belang is om verhalen te voorzien van verbindingen. Verbindingen maken de relaties van verhalen duidelijk: tussen verhalen, tussen verhalen en objecten in collecties en tussen verhalen en onderwerpen en entiteiten waarover de verhalen gaan. Verbindingen zorgen ervoor dat verhalen beter gevonden kunnen worden door gebruikers.

Verbindingen zijn onderdeel van de structuur van een verhaal, zoals besproken in het voorgaande hoofdstuk. Een verhaal kan verbindingen krijgen door er termen aan toe te kennen. Een term is een beschrijving die een onderwerp of entiteit uniek identificeert en de betekenis ervan duidelijk maakt. Een term wordt beheerd in een terminologiebron, zoals Art & Architecture Thesaurus, GeoNames of RKDartists.

Het wordt steeds eenvoudiger voor instellingen om terminologiebronnen te gebruiken. Een bekend hulpmiddel is het Termennetwerk van het NDE⁸. Verhalenmakers kunnen deze zoekmachine raadplegen om termen te vinden die hun verhalen het beste duiden. Desondanks komt er veel handwerk kijken bij het gebruik van het Termennetwerk, met als gevolg dat verhalen mogelijk te weinig verbindingen krijgen. Dat willen wij oplossen.

Software

Wij willen het voor verhalenmakers makkelijker maken om termen toe te kennen aan verhalen. Een manier om dat te doen is met software. Daarom hebben wij een kleine, experimentele applicatie ontwikkeld: SNEL, *Simple Named Entity Linking*⁹. De applicatie werkt als volgt:

1. Een verhalenmaker schrijft een verhaal in zijn of haar verhalenbeheersysteem, zoals een collectieregistratie- of een contentmanagementsysteem. Vervolgens geeft de maker aan in welke terminologiebronnen termen voorkomen waarover het verhaal gaat. Bijvoorbeeld: als een verhaal kerken in Noord-Brabant als onderwerp heeft, selecteert de maker de terminologiebron 'Brabantse gebouwen'; als een verhaal vertelt over plaatsen in Limburg, selecteert de maker de terminologiebron 'GeoNames'.

⁸ <https://termennetwerk.netwerkdigitaalnerfgoed.nl>

⁹ Broncode: <https://github.com/sdevalk/simple-named-entity-linking/>

2. Het verhalenbeheersysteem verstuurt de tekst van het verhaal en de gekozen terminologiebronnen naar SNEL. SNEL analyseert de tekst en haalt er de zogeheten *named entities*¹⁰ uit, zoals namen van locaties of personen. Vervolgens zoekt SNEL voor elke naam in de *reconciliation service*¹¹ van de terminologiebronnen. De reconciliation service geeft daarop de termen terug die overeenkomen met de naam. Bijvoorbeeld: de naam 'Mariadal' levert de term 'Klooster Mariadal' op in 'Brabantse gebouwen'; de naam 'Sittard-Geleen' levert de term 'Gemeente Sittard-Geleen' op in 'GeoNames'.
3. De verhalenmaker krijgt in zijn of haar verhalenbeheersysteem daarna de gevonden termen te zien. De maker kan vervolgens aangeven welke termen daadwerkelijk relevant zijn voor het verhaal en die bij het verhaal opslaan.

De aanpak van SNEL heet *named entity linking*: het linken van termen uit terminologiebronnen aan *named entities*¹². Voor zover wij weten is SNEL de eerste applicatie die aan *named entity linking* doet met behulp van *reconciliation services*, waaronder die van het Termennetwerk.

SNEL is bedoeld om in verhalenbeheersystemen gebruikt te worden, via een zogenaamde *Application Programming Interface* (API). Om de mogelijkheden van de applicatie te kunnen laten zien, hebben wij ook een *demonstrator* gemaakt¹³. Dit is een eenvoudige webpagina waar verhalenmakers de tekst van hun verhaal kunnen plakken en vervolgens de termen in de tekst kunnen laten opzoeken door SNEL. Een screenshot van de demonstrator staat op de volgende bladzijde.

Helpt SNEL om termen toe te kennen aan verhalen? Ons antwoord op basis van het experiment: ja, dat is het geval, maar met deze belangrijke aandachtspunten:

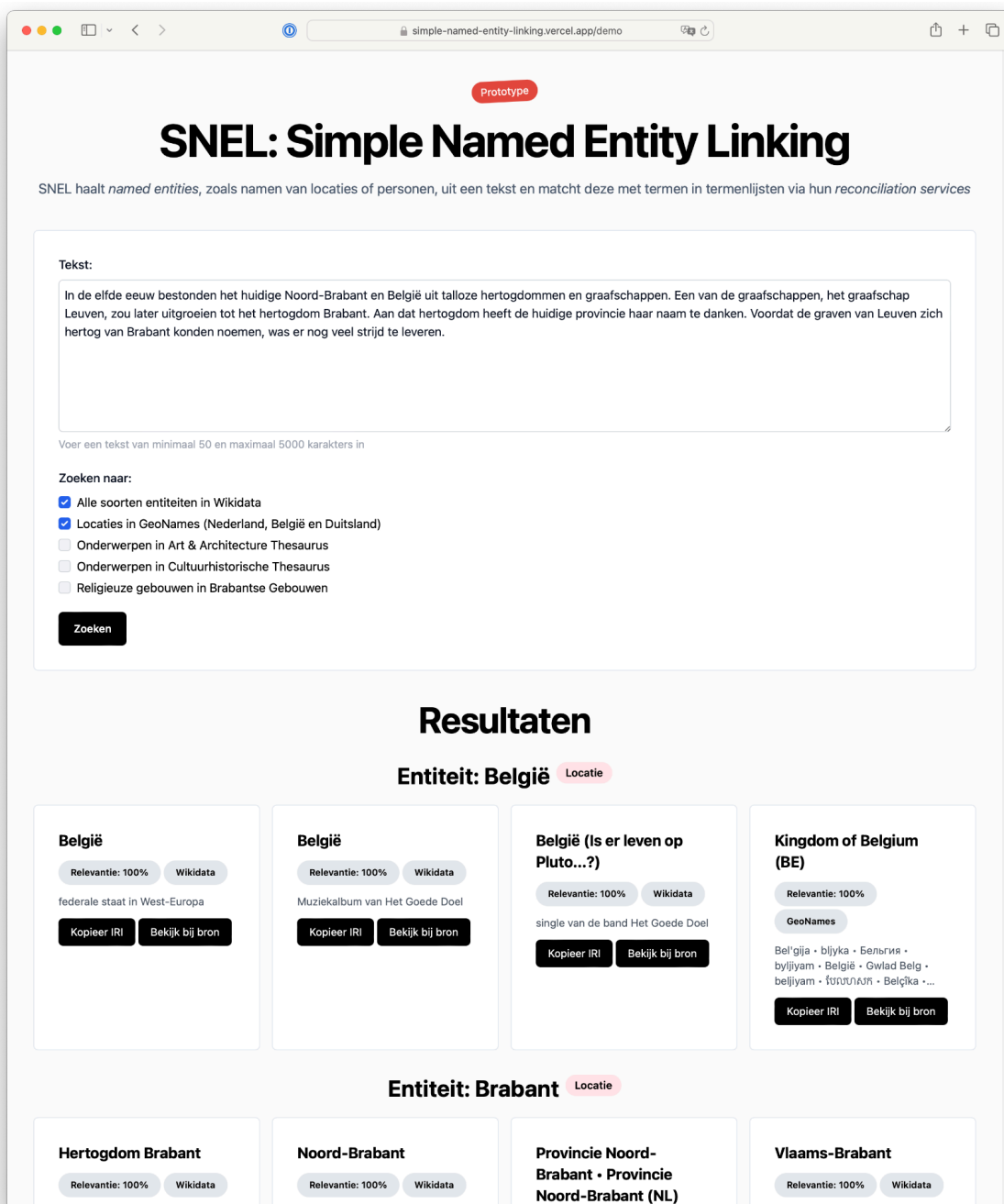
1. SNEL zoekt nu naar termen voor *alle* *named entities* in een verhaal, maar niet elke entity is even belangrijk. Bijvoorbeeld: als een verhaal over klooster Mariadal gaat maar terloops Groningen noemt, dan toont SNEL de termen die verband houden met deze plaats ook al gaat het verhaal hier niet over. Een oplossing hiervoor is door enkel termen op te zoeken van *named entities* die echt van belang zijn voor het verhaal.
2. SNEL houdt nu geen rekening met de context van een verhaal. Bijvoorbeeld: als een verhaal gaat over het ontstaan van de provincie Noord-Brabant en meldt dat het voortkomt uit het graafschap Leuven in het huidige België, dan toont SNEL de term van de staat België en de term van de gelijknamige single van de band Het Goede Doel. De eerste term is mogelijk relevant, de tweede beslist niet. Een oplossing hiervoor is door enkel termen op te zoeken die gerelateerd zijn aan het verhaal.
3. SNEL vindt soms veel termen bij een verhaal. De resultaten moeten dan ook op een toegankelijke manier getoond worden zodat verhalenmakers eenvoudig en snel de belangrijkste termen kunnen selecteren. Dit doet een beroep op de presentatiemogelijkheden van hun verhalenbeheersystemen.

¹⁰ https://en.wikipedia.org/wiki/Named_entity

¹¹ <https://termennetwerk.netwerkdigitaalerfgoed.nl/reconciliation>

¹² https://en.wikipedia.org/wiki/Entity_linking

¹³ <https://simple-named-entity-linking.vercel.app> (tijdelijke locatie)



Prototype

SNEL: Simple Named Entity Linking

SNEL haalt *named entities*, zoals namen van locaties of personen, uit een tekst en matcht deze met termen in termenlijsten via hun *reconciliation services*

Tekst:

In de elfde eeuw bestonden het huidige Noord-Brabant en België uit talloze hertogdommen en graafschappen. Een van de graafschappen, het graafschap Leuven, zou later uitgroeien tot het hertogdom Brabant. Aan dat hertogdom heeft de huidige provincie haar naam te danken. Voordat de graven van Leuven zich hertog van Brabant konden noemen, was er nog veel strijd te leveren.

Voer een tekst van minimaal 50 en maximaal 5000 karakters in

Zoeken naar:

- Alle soorten entiteiten in Wikidata
- Locaties in GeoNames (Nederland, België en Duitsland)
- Onderwerpen in Art & Architecture Thesaurus
- Onderwerpen in Cultuurhistorische Thesaurus
- Religieuze gebouwen in Brabantse Gebouwen

Zoeken

Resultaten

Entiteit: België Locatie

België

Relevantie: 100% Wikidata

federale staat in West-Europa

Kopieer IRI **Bekijk bij bron**

België

Relevantie: 100% Wikidata

Muziekalbum van Het Goede Doel

Kopieer IRI **Bekijk bij bron**

België (Is er leven op Pluto...?)

Relevantie: 100% Wikidata

single van de band Het Goede Doel

Kopieer IRI **Bekijk bij bron**

Kingdom of Belgium (BE)

Relevantie: 100% GeoNames

Bel'gija · blyjka · Бельгия · byljiyam · België · Gwlad Belg · beljiyam · ბელჯიამი · Belçika · ...

Kopieer IRI **Bekijk bij bron**

Entiteit: Brabant Locatie

Hertogdom Brabant

Relevantie: 100% Wikidata

Noord-Brabant

Relevantie: 100% Wikidata

Provincie Noord-Brabant · Provincie Noord-Brabant (NL)

Vlaams-Brabant

Relevantie: 100% Wikidata

Figuur 2: screenshot van de demonstrator van SNEL

5. Hoe verder?

Wij hebben een aantal waardevolle inzichten opgedaan door de uitvoering van dit korte onderzoek. Op basis hiervan gaan wij een aantal acties ondernemen, zoals:

1. Bepalen hoe wij de duurzame toegankelijkheid van verhalen kunnen bewerkstelligen zodat ze niet alleen nu maar ook later gebruikt kunnen worden.
2. Inventariseren van de eisen – of *requirements* – die wij aan de applicaties stellen waarmee verhalen beheerd en gepubliceerd kunnen worden.
3. Inzichtelijk maken welk linked data-model wij kunnen hanteren voor het ontsluiten van verhalen, zodat verhalen benut kunnen worden in toepassingen van onszelf en anderen. Bijvoorbeeld: mogelijk kunnen Schema.org (een eenvoudig model) en/of CIDOC-CRM (een rijk model) benut worden.
4. Inventariseren welke systemen wij kunnen gebruiken voor het beheren en publiceren van verhalen. Wellicht kunnen wij hiervoor bestaande systemen inrichten, misschien is een nieuw type systeem nodig.
5. Bepalen hoe de software voor het geautomatiseerd toekennen van termen aan verhalen verbeterd en productiewaardig gemaakt kan worden.

Colofon

Dit is een uitgave van Erfgoed Brabant, in opdracht van het Netwerk Digitaal Erfgoed en gemaakt door:

Wouter Daemen (adviseur digitalisering, Huis voor de Kunsten Limburg)

Cris Kremers (teamleider Brabant Cloud, Erfgoed Brabant)

Sjors de Valk (zelfstandig IT-architect en software developer)

Deze uitgave kwam tot stand met dank aan de input van verschillende medewerkers van erfgoedinstellingen tijdens bijeenkomsten in 's-Hertogenbosch en Roermond.

Reacties zijn welkom via criskremers@erfgoedbrabant.nl.

