

Verbinden van collecties

Een verkenning

Datum: februari 2023
Auteur: Mirjam Verloop



podium
kunst
.net



Inhoudsopgave

1. Inleiding	3
2. Situatieschets	3
2.1 Belang	3
2.2 Overlap	3
2.2.1 Disciplines	3
2.2.2 Producenten, gezelschappen, podia en personen	3
2.2.3 Producties	4
3. Aanbevelingen	4
3.1 Terminologiebronnen	4
3.1.1 Personen en Instellingen	4
3.2 LRM/RDA en Linked Data	5
3.3 Verschillen duiden en afspraken maken	6
3.3.1 Definities	6
3.3.2 Basisgegevens	7
3.4 Overige aandachtspunten	7
4. Geraadpleegde website en informatie	8

1. Inleiding

Het **einddoel van Podiumkunst.net** is een integraal en toegankelijk digitaal overzicht van het Nederlandse podiumkunst-erfgoed. Een overzicht dat de geschiedenis van de Nederlandse podiumkunsten levend maakt en een inspiratiebron voor makers van nu en de toekomst kan vormen.¹

In 2022 voerde Podiumkunst.net een verkenning uit naar de overlap én de verschillen tussen twee collecties van twee verschillende instellingen. De ene collectie bevat veelal **historische** gegevens, de andere collectie meer **actuele** gegevens. Het doel van de verkenning was om inzicht te krijgen in wat nodig is om deze collecties met elkaar te verbinden. De verkenning heeft inzichten opgeleverd die bruikbaar zijn voor alle collecties in de podiumkunstensector.

Tijdens deze verkenning is meermaals gesproken met vertegenwoordigers van beide instellingen. Daarnaast zijn beide collecties in detail onderzocht en is aanvullende documentatie gelezen.

2. Situatieschets

2.1 Belang

Vertegenwoordigers van beide instellingen leggen de nadruk op het belang van verbonden collecties en archivering van informatie voor de hele sector. Een koppeling - en publicatie van bijvoorbeeld termenlijsten - opent bovendien mogelijkheden om te verrijken. Tot slot opent de koppeling mogelijkheden om te delen met andere organisaties zoals de Theaterkrant en toneelbibliotheken, en het hergebruiken van gegevens in de sector.

2.2 Overlap

De overlap van de verzamelde gegevens in beide collecties is te vinden bij disciplines, personen en instellingen (producenten, gezelschappen, podia) en producties.

2.2.1 Disciplines

Een discipline duidt het soort productie, bijvoorbeeld muziektheater, opera of kleinkunst. De gebruikte disciplines komen deels overeen voor beide collecties.

2.2.2 Producenten, gezelschappen, podia en personen

In beide collecties is informatie te vinden over podia, gezelschappen en producenten. Namen en adresgegevens komen in beide collecties voor. De collectie met historische gegevens heeft ook een uitgebreid overzicht van personen.

¹ [Activiteitenplannen - Podiumkunst.net](#)

2.2.3 Producties

Beide collecties bewaren informatie over producties en uitvoeringen. Een aantal velden overlapt in meer of mindere mate. Het gaat om titel, producent, de eerdergenoemde discipline, datum, zaal/locatie, festival.

Er zijn nuanceverschillen. Zo bewaart de collectie met historische gegevens naast de titel ook de oorspronkelijke titel. Dit komt met name voor bij producties in een andere taal, bijvoorbeeld een Friese productie.

Ook de seizoenswissel is voor beide instellingen anders. Voor de collectie met meer actuele gegevens wisselt het seizoen op 1 juli. Het lopende seizoen is 2022, het vorige seizoen 2021 en het volgende seizoen 2023. Voor de collectie met historische gegevens wisselt het seizoen op 1 augustus. De duiding van het seizoen is in die collectie in de vorm van een trefwoord, bijvoorbeeld *Seizoen 1988/89*.

3. Aanbevelingen

3.1 Terminologiebronnen

In de gesprekken met beide instellingen komen termenlijsten of terminologiebronnen regelmatig ter sprake. Er is behoefte aan eenduidigheid. Beide collecties werken met interne bronnen of lijsten voor bijvoorbeeld personen, instellingen of disciplines.

Termen beschrijven waar erfoegd over gaat. Toch is een term meer dan een woord. Elke term heeft namelijk een unieke identifier, een URI. Een URI is een uniek adres waardoor ondubbelzinnig duidelijk is welke term bedoeld wordt.²

3.1.1 Personen en Instellingen

Een terminologiebron voor producenten, gezelschappen, personen en podia lijkt voor de hand te liggen. Het maakt zonder twijfel duidelijk om welke persoon of instelling het gaat en het is een manier om gegevens te verbinden. Het voorkomt bovendien een mix van (onduidelijke) schrijfwijzen. Bovendien biedt het mogelijkheid om *equivalenten*, *broader* en *narrower terms* te definiëren.

Voorbeelden van verschillende schrijfwijzen:

- *Nationale Theater, Het en Het Nationale Theater*
- *Kan, Wim en Wim Kan*

² [Veelgestelde vragen over het Termennetwerk](#)

Het samenstellen en publiceren van een terminologiebron met personen, gezelschappen en podia lijkt een waardevolle stap. Het maakt koppeling mogelijk en voorkomt verwarring in de eigen systemen, in de sector en bij eindgebruikers van gegevens.

Suggesties

1. Verdere verkenning: Wat is de beste aanpak om tot een bron met personen en instellingen te komen?
2. Stem de aanpak af met de werkgroep metadata van Podiumkunst.net.
3. Publiceer de terminologiebron via het [Termennetwerk NDE](#) en maak een handleiding voor het gebruik van de bron.
4. Breng de bron onder de aandacht bij instellingen in de sector en motiveer het gebruik van de bron.

Aandachtspunten

1. Zijn er aandachtspunten met betrekking tot privacy van (persoons)gegevens?
2. Wie wordt de kartrekker en eigenaar van deze bron? Hoe en wie houdt de lijst actueel?
3. Welke consequenties heeft dit voor huidige data van instellingen? Welke inspanning is nodig om bestaande gegevens te vervangen door termen of om termen toe te voegen?
4. Zijn er bestaande, openbare personenregisters die hier bruikbaar zijn? Bijvoorbeeld via [IISG](#).

3.2 LRM/RDA en Linked Data

Het conceptuele deel voor de Podiumkunst.net infrastructuur, de *knowledge graph*, is gebaseerd op het LRM/RDA-model. De kern van dit model is de zogenaamde *WEMI*-structuur: Werken, Expressies, Manifestaties en Items. Hieraan zijn diverse andere entiteiten gerelateerd (zoals Personen, Plaatsen, Periodes en - meer generiek - 'Dingen' (Res) en aanduidingen (Nomen)).³

De werkgroep metadata van Podiumkunst.net heeft in 2021 na onderzoek gekozen voor dit RDA model om gegevens te beschrijven. Deze internationale standaard maakt metadata geschikt om te publiceren als Linked Open Data (LOD). Uit het onderzoek bleek dat RDA het best aansluit bij de al gecreëerde metadata in de sector. RDA helpt om de rijkdom in collecties

³ [De Podiumkunst.net knowledge graph: Het hart van de digitale infrastructuur die podiumkunst-collecties samenbrengt](#)

te vinden en verbanden zichtbaar te maken. Vanwege de gedetailleerdheid is RDA goed geschikt voor ontsluiting van collectiegegevens.

Het RDA model helpt om collecties op een eenduidige manier te beschrijven. Een collectie die volledig *gemapt* is aan het RDA-model kan uiteindelijk worden gepubliceerd als Linked Data.

Suggesties

1. Verdere verkenning: wat betekent het voor instellingen en hun collecties om te conformeren aan RDA? Op welke manier kan *gemapt* worden?
2. Podiumkunst.net heeft kennis opgedaan in andere projecten.⁴ Haal die kennis op, borg en gebruik in toekomstige projecten.
3. Maak afspraken over het uniform gebruik van RDA, bijvoorbeeld in een applicatieprofiel.
4. Vergroot de kennis van RDA door trainingen voor verschillende kennisniveaus..
5. Ondersteun instellingen bij het *mappen* van collecties naar RDA.

3.3 Verschillen duiden en afspraken maken

Naast overlappende gegevens zijn er ook verschillen. Soms klein, soms groter. Dit hoeft niet altijd erg te zijn, toch is het goed om deze verschillen te duiden en er eventueel (sectorbrede) afspraken over te maken om verwarring te voorkomen.

3.3.1 Definities

In vrijwel alle gesprekken is gerefereerd aan definities, en dan met name:

- *Wat is een productie?*
- *Wie is de eigenaar van een productie?*
- *Wat is een uitvoering?*
- *Wie is de eigenaar van een uitvoering?*
- *Wanneer is een productie (of première) een nieuwe productie (première)?*
- *Van wanneer tot wanneer loopt een seizoen?*
- *Hoe noemen we een seizoen?*

Een concreet verschil is de eerdergenoemde seizoensduiding. De ene instelling wisselt van seizoen op 1 juli, de andere instelling wisselt van seizoen op 1 augustus.

⁴ [De afronding van het project Musidesk ouverture als blauwdruk voor andere organisaties - Podiumkunst.net](#)

3.3.2 Basisgegevens

Op basis van de inventarisatie van de collecties is te duiden welke gegevens overlappen. Deze basisgegevens lijken nodig om te verbinden. Het gaat dan om: Titel van de voorstelling, Producent/gezelschap, Discipline, Voorstellingsdatum, Naam van producent, gezelschap of theater en (beperkte) Adresgegevens van de producent, gezelschap of theater.

Ook voor deze gegevens geldt dat er soms kleine afwijkingen zijn in notatie, zoals *Het Nationale Theater* en *Nationale Theater, Het*. Deze verschillen kunnen grotendeels opgelost worden door het gebruik van eenduidige bronnen en het maken van afspraken.

Suggesties

1. Zorg voor heldere definities in de sector, dit vergroot de kans op een geslaagde uitwisseling van gegevens.
2. Maak een goede lijst van basisgegevens ('wat is er minimaal nodig om te verbinden?') en maak afspraken over de nuanceverschillen. Deze lijst van basisgegevens vormt een vertrekpunt voor het verbinden.
3. Gebruik de expertise van de metadata werkgroep.

3.4 Overige aandachtspunten

Tot slot, nog een aantal aandachtspunten vanuit de gesprekken en het onderzoek.

3.4.1 Privacy en auteursrecht

Wanneer gegevens in een terminologiebron of dataset worden opgenomen of beschikbaar komen via een koppeling, dan is het goed om te toetsen of al die gegevens gedeeld mogen worden vanuit privacy en auteursrechtelijke status.

Het is aannemelijk dat adresgegevens van een instelling, bijvoorbeeld een podium, openbaar zijn en gedeeld kunnen worden. Voor adresgegevens van personen is dat anders.

Consequenties voor instellingen

Instellingen kunnen de consequenties nog niet goed overzien van bijvoorbeeld de overstap naar het publiceren in Linked Data conform RDA of het gebruik van terminologiebronnen. Ze vragen zich bijvoorbeeld af hoeveel tijd dat kost en wat de gevolgen zijn voor de historische gegevens.

Kennis en vaardigheden

Op sommige vlakken, zoals RDA en Linked Open Data mist kennis bij de instellingen. Training en begeleiding kan instellingen helpen.

4. Geraadpleegde website en informatie

Bram Gaakeer, Remco de Boer, Enno Meijers, Sjors de Valk, Marcel Ras, Tamara van Zwol, Joost van der Nat, Annelot Vijn, Willem Melder, Erik van den Bergh, Steven Ham, Frans van der Zande, Laurents Sesink, Marcus Cohen, Ivo Dahlmans, Michelle Boon, Gijs Broos & Mark van Hoorn. (2021). Digitaal Erfgoed Referentie Architectuur (DERA) - versie 4.0 (4.0). Zenodo.
<https://doi.org/10.5281/zenodo.5562062>

Cultuurindex Boekmanstichting (z.d.) Theater: kijkend naar de toeschouwer.
Boekmanstichting. Kenniscentrum voor kunst, cultuur en beleid.
[Theater - Boekmanstichting](#)

Datasetregister (z.d.). Veelgestelde vragen over het Datasetregister. Netwerk Digitaal Erfgoed.
[Datasetregister - Voor alle erfgoeddatasets!](#)

Podiumkunst.net (2022, 7 april). De Podiumkunst.net knowledge graph: Het hart van de digitale infrastructuur die podiumkunst-collecties samenbrengt. Podiumkunst.net.
[De Podiumkunst.net knowledge graph: Het hart van de digitale infrastructuur die podiumkunst-collecties samenbrengt](#)

Podiumkunst.net (2021, 18 augustus). Het verbinden en ontsluiten van collecties uit de podiumkunsten: metadata en het RDA-model. Podiumkunst.net. [Het verbinden en ontsluiten van collecties uit de podiumkunsten: metadata en het RDA-model](#)

Termennetwerk (z.d.) FAQ. Netwerk Digitaal Erfgoed.
<https://termennetwerk.netwerkdigitaal erfgoed.nl/faq>

podium
kunst
.net

podium
kunst
.net

podium
kunst
.net

podium
kunst
.net

podium
kunst
.net

podium
kunst
.net

podium
kunst
.net

p
k
.n

odium
unst
et

podium
kunst
.net