

Verwerking van studiedata aan De HHs

Beschrijving en verantwoording, v. 0.9.0



let's change
YOU. US. THE WORLD.

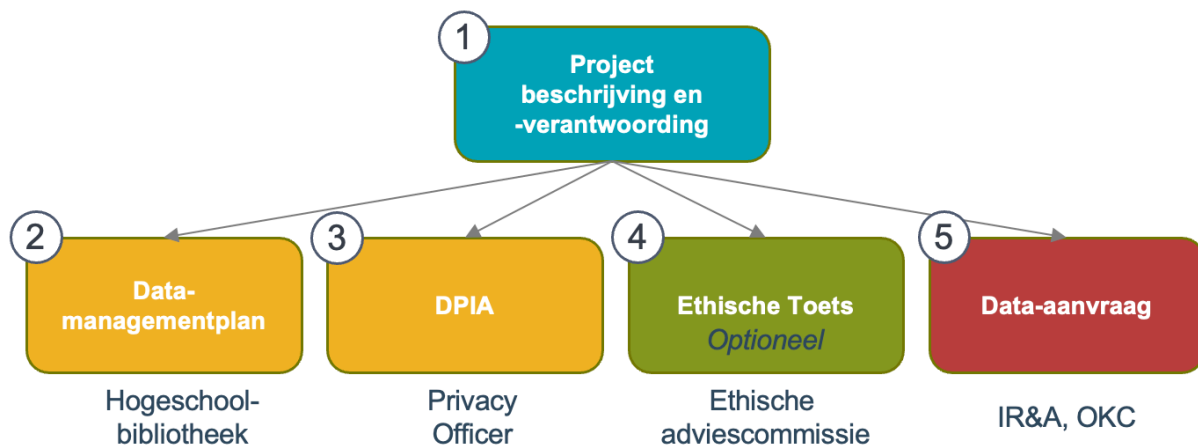
Inhoudsopgave

1	Inleiding	2
1.1	Verwerking van studiedata aan De HHs	2
1.2	Doel van deze handleiding	2
2	Procedures voor levering van studiedata	3
2.1	Levering aan het lectoraat	3
2.2	Projectbeschrijving	3
2.3	Datamanagementplan	4
2.4	Data Privacy Impact Assessment (DPIA)	4
2.4.1	Keuze voor een grondslag	5
2.4.2	Grondslag: Gerechtvaardigd belang	5
2.4.3	Grondslag: Toestemming	6
2.5	Toets bij de Ethische Adviescommissie	6
2.6	Data-aanvraag bij IR&A, OKC	7

1 Inleiding

1.1 Verwerking van studiedata aan De HHs

In het lectoraat Learning Technology & Analytics onderzoeken we studiedata van De HHs. Hiervoor verwerken we studiedata die door de dienst Onderwijs, Kennis- en Communicatie (OKC) vanuit verschillende informatiesystemen gecombineerd en geleverd kunnen worden. Voor de levering aan het lectoraat dient een aantal formele stappen doorlopen te worden.



1.2 Doel van deze handleiding

Deze handleiding van het lectoraat Learning Technology & Analytics beschrijft en verantwoordt de procedures en methode om deze studiedata aan te vragen en te verwerken. Bij subsidieaanvragen dienen deze documenten meestal meegeleverd te worden.

Heb je suggesties voor verbeteringen? Mail die dan naar Theo Bakker, t.c.bakker@hhs.nl.

2 Procedures voor levering van studiedata

2.1 Levering aan het lectoraat

Voor de levering van studiedata dienen een aantal procedures doorlopen te worden en bijpassende documenten opgesteld te worden:

1. Projectbeschrijving en -verantwoording
2. Datamanagementplan
3. DPIA
4. Eventueel een toets bij de Ethische Adviescommissie
5. Formele data-aanvraag bij het team IR & Analytics, OKC.

! Belangrijk

Stem voordat je een of meerdere van deze documenten formeel instuurt, deze eerst inhoudelijk af met de lector.

2.2 Projectbeschrijving

De projectbeschrijving bevat:

1. Een **toelichting op het belang van het onderzoek**: voor studenten en docenten of andere stakeholders in onze hogeschool, de doelstellingen van De HHs, de onderzoekslijnen van het lectoraat, de onderzoeksthema's van het kenniscentrum Global & Inclusive Learning, en de onderzoeksthema's van De HHs.
2. Een **theoretische en ethische onderbouwing**. De theoretische onderbouwing bevat een link met het vakgebied en eerdere publicaties; de ethische onderbouwing maakt gebruik van de [Data Science Ethics Checklist](#) van Deon die ingevuld als bijlage wordt toegevoegd.
3. Een uitwerking van de **onderzoeksvragen en operationalisering** daarvan.
4. Een uitwerking van de **databronnen**.
5. Een beschrijving van de **levering en verwerking**.
6. De **methoden van analyse**.
7. De **verwachte resultaten**.
8. Een **toelichting op de reproduceerbaarheid** van het onderzoek volgens de [FAIR](#) principes.
9. Bijlagen
 1. **Referenties**
 2. Een **versiegeschiedenis** met wijzigingen en verspreiding.
 3. De locatie van de **repository** voor de projectbeschrijving.
 4. De ingevulde Data Science Ethics Checklist

Laat de projectbeschrijving lezen door een aantal collega's zodat deze aan kwaliteit wint. Geef bij een verzoek voor een review expliciet aan in het verzoek dat het een formeel verzoek is, wanneer je een reactie hoopt te ontvangen en hoe je de feedback zult gaan gebruiken.

Template: Gebruik voor de projectbeschrijving het markdown template van het lectoraat.

2.3 Datamanagementplan

De hogeschoolbibliotheek van de dienst OKC heeft datamanagement in de portefeuille. Voor het datamanagementplan gebruiken we een standaard formulier van NWO. Het datamanagementplan beschrijft:

1. **Algemene gegevens:** informatie over het project, de aanvrager, de datasteward
2. **De onderzoeksdata:** welke data worden verzameld of geproduceerd, en welke bestaande data worden hergebruikt?
3. **Metadata:** welke metadata en documentatie worden meegeleverd met de data?
4. **Opslag en back-up:** Hoe worden data en metadata opgeslagen en geback-up't tijdens het onderzoek?
5. **Bewaartermijnen en archivering:** Hoe en wanneer worden data voor langere tijd gedeeld en bewaard?
6. **Datamanagementkosten**

Omdat het hier om de beveiliging van data gaat, is het goed het datamanagementplan voorafgaand aan de DPIA op te stellen aangezien dit type vragen ook terugkomt in de DPIA.

Je bespreekt het datamanagementplan met een medewerker van de hogeschoolbibliotheek met wie je een afspraak kan inplannen via research@hhs.nl.

Template: Gebruik voor het datamanagementplan het standaard NWO formulier.

2.4 Data Privacy Impact Assessment (DPIA)

De inschrijvingen of resultaten dataset kan na positieve afronding van de DPIA geanonimiseerd aanleverd worden door OKC. Herkenbare gegevens, zoals voor- en achternaam, geboortedatum, adres zijn verwijderd en het studentnummer is vervangen door een willekeurige tekst, waarop beide sets wel gekoppeld kunnen worden.

Geanonimiseerde gegevens vallen formeel niet onder de Algemene Verordening Gegevensbescherming (AVG). Maar omdat datasets met veel gegevens mogelijk toch herleidbaar is tot een student, spreken we liever van gepseudonimiseerde data en voeren we toch een Data Privacy Impact Assessment (DPIA) uit.

Template: De DPIA wordt uitgevoerd aan de hand van de DPIA vragenlijst van De HHs.

2.4.1 Keuze voor een grondslag

Lees voor meer informatie over privacy, ethiek en studiedata het [Referentiekader privacy en ethiek voor studiedata](#).

De keuze voor de grondslag voor de verwerking van studiedata voor het onderzoek en de onderbouwing daarvan zijn cruciaal. De grondslag voor de verwerking door het lectoraat is het **gerechtvaardigd belang** van het onderzoek voor De Haagse Hogeschool of – uitsluitend bij onderzoek met proefpersonen of bijzondere persoonsgegevens¹ – de **toestemming** van studenten. Zodra je op basis van data een advies geeft aan studenten heb je toestemming nodig van de student om zijn/haar/diens persoonsgegevens te verwerken.

2.4.2 Grondslag: Gerechtvaardigd belang

Voor de grondslag gerechtvaardigd belang zijn [3 voorwaarden](#):

Voorwaarde 1: gerechtvaardigd belang

Uw belang is echt een gerechtvaardigd belang als het ergens in het recht is opgenomen. En wordt erkend en beschermd. Dat mag ook in een ongeschreven rechtsregel of rechtsbeginsel zijn. Als het maar gaat om een belang waarvan we in de maatschappij vinden dat het door het recht beschermd moet worden.

Hier geldt het belang van wetenschappelijk onderzoek naar onderwijs in het algemeen en het onderwijs van De Haagse Hogeschool in het bijzonder.

Voorwaarde 2: noodzakelijkheid

Heeft u daadwerkelijk een gerechtvaardigd belang? Dan moet u vervolgens kijken of de verwerking van persoonsgegevens noodzakelijk is om dit belang te behartigen. Dit doet u door na te gaan:

- Of het doel van uw verwerking in verhouding staat tot de inbreuk op de privacy van de betrokkenen. In de AVG heet dit ‘proportionaliteit’.
- Of u het doel niet op een andere manier kunt bereiken, die minder ingrijpend is voor de betrokkenen. In de AVG heet dit ‘subsidiariteit’.
- *Proportionaliteit* - Voor ons onderzoek gebruiken we gepseudonimiseerde data, waarbij we zo min mogelijk gegevens verzamelen. Bovendien proberen we met behulp van statistische analyses en machine learning het aantal variabelen te reduceren tot die variabelen die significant zijn of substantieel bijdragen aan accuraatheid van voorspellingen.

¹ [Bijzondere persoonsgegevens](#) zijn iemands gegevens over ras of etniciteit, politieke opvatting, religieuze of levensbeschouwelijke overtuigingen, lidmaatschap van een vakvereniging, gezondheid, seksueel gedrag of seksuele gerichtheid, genetische gegevens en biometrische gegevens voor unieke identificatie.

- *Subsidiariteit* - Er is geen mogelijkheid deze gegevens op een andere manier te verzamelen.

Voorwaarde 3: afweging belangen

Heeft u een gerechtvaardigd belang en is de gegevensverwerking noodzakelijk om dit belang te behartigen? Dan moet u tot slot een afweging maken tussen uw belangen en de belangen van de betrokkenen.

Bij deze afweging kijkt u naar:

- de gevolgen voor de betrokkenen;
- hoe ernstig de inbreuk is op de privacy van de betrokkenen; welke (aanvullende) maatregelen u heeft genomen om ongewenste gevolgen voor de betrokkenen te voorkomen of beperken;
- of de betrokkenen de verwerking min of meer kunnen verwachten. Bijvoorbeeld als vervolg op een eerdere verwerking waarvoor zij toestemming hebben gegeven of als vervolg op verwerkingen die noodzakelijk zijn om een contract uit te voeren.

In de projectbeschrijving wordt de afweging van belangen beschreven.

2.4.3 Grondslag: Toestemming

Toestemming wordt ook wel *informed consent* genoemd. Hierbij geeft de student akkoord op de verwerking van diens persoonsgegevens. Dit is van toepassing bij bijzondere persoonsgegevens of bij profiling, waarbij de student op basis van persoonsgegevens een advies wordt gegeven of een interventie krijgt aangeboden. Denk aan deelname aan een gericht experiment, bijv. de toepassing van een algoritme voor een prognose.

2.5 Toets bij de Ethische Adviescommissie

Een ethische toets is altijd van belang voor een project. Dit neem je op in je projectbeschrijving; zie voor meer info . Als je een onderzoek doet met proefpersonen of er een risico bestaat op negatieve gevolgen van je onderzoek, dan is het van belang dit voor te leggen aan de [Ethische Adviescommissie](#) van De HHs.

Voor de toets is er een checklist opgesteld die je ingevuld opstuurt samen met de projectbeschrijving. De checklist van de commissie bevat 1) vragen over de aanvrager, 2) de ethische adviesvragen, 3) het onderzoeksproject in het algemeen, 4) de uitvoering en methoden van het onderzoek, 5) de deelnemers/respondenten en 6) eventueel aanvullende vragen of opmerkingen.

De commissie vergadert 1 keer per maand. De data vind je op de webpagina van de commissie. Je stuurt je ingevulde checklist naar research@hhs.nl ter attentie van de ethische commissie.

De Ethische Adviescommissie adviseert jou als onderzoeker over ethische aspecten ten aanzien van jouw onderzoeksproject. Het advies vindt bij voorkeur vóór of bij de start van jouw onderzoeksactiviteiten plaats, maar advies kan indien nodig ook gedurende het onderzoekstraject worden ingewonnen (bijvoorbeeld wanneer een journal je vraagt om een ethisch adviesrapport).

De volgende vragen staan centraal bij de commissie:

- Hoe worden de belangen van de deelnemers aan jouw onderzoek beschermd?
- Worden deelnemers in jouw onderzoek voldoende en juist geïnformeerd over deelname?

Template: Gebruik voor een aanvraag aan de Ethische Adviescommissie de Checklist aanvraag Ethische Adviescommissie.

2.6 Data-aanvraag bij IR&A, OKC

Vanuit OKC is een **analyseset** gebouwd op basis van een aantal bronnen: Osiris, Studielink, het 1CHO bestand en een reistijden dataset. Deze dataset bevat data over alle inschrijvingen van studenten van De HHs vanaf cohort 2012. De methode voor het inlezen, verwerken en combineren van data is ontwikkeld door de Vrije Universiteit (VU) en overgenomen door De HHs. Zie voor een beschrijving van de variabelen *bijlage x*.

De data voor je onderzoek vraag je parallel aan bij het team IR & Analytics van OKC. Zodra er een akkoord is vanuit de privacy officer kan de data aan je geleverd worden. Een aanvraag doe je via Topdesk: [BIV: Rapportages over studenten en onderwijs](#) > tegel 'Nieuw overzicht aanvragen'.

Het formulier vraagt naar 1) algemene gegevens over de aanvrager, 2) of de gegevens die je aanvraagt persoonlijke gegevens bevat, 3) algemene informatie over je aanvraag: omschrijving, doel, welke gegevens je nodig hebt, periode, selectie, overige randvoorwaarden/wensen, wie er toegang krijgt tot de data, wanneer je de dataset nodig hebt.

Bij vraag 2 is er een optie voor Ja of Nee, ze zijn geanonimiseerd en geaggregeerd. Dit zijn antwoord opties die niet goed aansluiten bij de gepseudonimiseerde data op studentniveau die we in het lectoraat onderzoeken. Ik heb het IR&A team gevraagd dit aan te passen. Kies voor nu de 2e optie met een toelichting.

Niet alle data zijn al beschikbaar: enkel inschrijvingen en data over uitval of diplomering van de cohorten 2012 tot en met 2022. Data over de faculteit ITD ontbreekt nog. Een resultaatendataset is in voorbereiding en gepland om voor de zomer af te ronden. Data uit Brightspace of andere systemen is nog niet ontsloten.

Template: Gebruik het formele aanvraagformulier