



**Centraal Bureau
voor de Statistiek**

Microdata Services

Documentatie

**Kenmerken patentaanvragers en patentaanvragen
naar activiteit, grootteklasse, land van zeggenschap
en technologiegebied
(PAT)**

Bronvermelding

Publicatie van uitkomsten geschiedt door de onderzoeksinstituting of de opdrachtgever op eigen titel. Verwijzing naar het Centraal Bureau voor de Statistiek (CBS) betreft uitsluitend het gebruik van de niet-openbare microdata. Deze microdata zijn onder [bepaalde voorwaarden](#) voor statistisch en wetenschappelijk onderzoek toegankelijk. Voor nadere informatie microdata@cbs.nl. Dat wordt als volgt geformuleerd:

“Resultaten [gedeeltelijk] gebaseerd op eigen berekeningen [*naam onderzoeksinstituting, c.q. opdrachtgever*] op basis van niet-openbare microdata van het Centraal Bureau voor de Statistiek betreffende *Kenmerken patentaanvragers en patentaanvragen naar activiteit, grootteklasse, land van zeggenschap en technologiegebied.*”

Engelse versie

“Results based on calculations by [*name of research institution or commissioning party*] using non-public microdata from Statistics Netherlands.”

Under [certain conditions](#), these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

Beschikbare bestand(en):

De bestanden zijn beschikbaar over de perioden 2013 t/m 2020.

In de [Versiegeschiedenis](#) ziet u een overzicht van de beschikbare bestanden.

De gebruiker dient rekening te houden met het volgende:

- Reden versie 2019V2 en 2020V2: De (nader) voorlopige cijfers over 2019 en 2020 zijn per 6 februari 2024 bijgewerkt.
 - Voor de persoonskenmerken en/of achtergronden dient u de beschikbare bestanden te raadplegen. Deze staan bij [Zelf onderzoek doen](#) in de catalogus onder het thema [Bevolking](#). Voor het aanvragen van deze bestanden geldt de gebruikelijke procedure.
 - Zie [Onderzoeksomschrijvingen](#) voor meer informatie over alle onderzoeken.
-



Inhoudsopgave

Bronvermelding.....	2
Beschikbare bestand(en):	3
1. Inleiding.....	5
2. Toelichting op de inhoud van de bestanden	6
Beschrijving	6
Beschrijving van de populatie	6
Methodologie	6
Procesverloop.....	6
Bijlagen	7
3. Bestandsopbouw en toelichting.....	8
Bestandsopbouw.....	8
Toelichting op de variabelen	9
Auth	9
Patstat-psn-id_crypt	9
Og-id	10
Docdb-family-id_crypt.....	10
Status_samw_btl	10
UCI.....	11
SBI_ABR	11
GK_SBS.....	11
Technologiegebied_IPC_CBS	12
Tech_weight_IPC_CBS	12
Technologiegebied_WIPO.....	13
Tech_weight_WIPO.....	13
Versiegeschiedenis.....	14

1. Inleiding

Het team Microdata Services van het CBS stelt, onder [bepaalde voorwaarden](#), niet-openbare microdata (gepseudonimiseerde data op persoons-, bedrijfs- en adresniveau) toegankelijk voor statistisch en wetenschappelijk onderzoek. Dit rapport beschrijft de inhoud en structuur van “*Kenmerken patentaanvragers en patentaanvragen naar activiteit, grootteklasse, land van zeggenschap en technologiegebied*”.

Hoofdstuk 2 beschrijft in het algemeen de microdatabestanden zoals populatieafbakening, methodologische bijzonderheden, kwaliteit en herkomst van de gegevens.

In hoofdstuk 3 wordt de bestandsindeling gegeven en worden voor de categoriale variabelen alle mogelijke scores en hun betekenis opgesomd.

2. Toelichting op de inhoud van de bestanden

Beschrijving

Deze dataset bevat gegevens over patentaanvragers- en aanvragen door Nederlandse ondernemingen bij het European Patent Office (EPO) en bij Octrooicentrum Nederland (OCNL), de patentverlenende instantie voor Nederland welke onderdeel is van de Rijksdienst voor Ondernemend Nederland (RVO). Patenten kunnen worden toegekend aan ondernemingen, universiteiten, instellingen of particulieren zolang de uitvinding voldoet aan de voorwaarden voor patenteerbaarheid, wat door een octrooibureau (EPO of OCNL) wordt getoetst. Om voor een patent in aanmerking te komen moet een uitvinding nieuw, inventief en ontvankelijk voor industriële toepassing zijn. Het geeft de eigenaar van het patent het wettelijke recht anderen te verbieden de uitvinding na te maken, te gebruiken, te verkopen of in te voeren. Het patent geldt voor een beperkte tijdsduur en alleen in die landen die vallen onder de bescherming van het patent.

Beschrijving van de populatie

De populatie betreft alle directe patentaanvragen uit Nederland bij het European Patent Office (EPO) en Octrooicentrum Nederland (OCNL), zoals verkregen uit de EPO Worldwide Patent Statistical Database (PATSTAT) online database welke tegen betaling beschikbaar wordt gesteld door de EPO. De gegevens die worden opgehaald uit de database betreffen uitsluitend informatie over Nederlandse aanvragers. PATSTAT is gebaseerd op de centrale EPO documentatie database (DOCDB). Het grootste deel van de patentaanvragers wordt gekoppeld aan een onderneming uit het Algemeen Bedrijvenregister (ABR).

Methodologie

Op basis van de naam-, adres- en woonplaatsgegevens van de aanvrager van het patent is via een algoritme een koppeling gelegd met de statistische eenheid onderneming uit het Algemeen Bedrijvenregister (ABR). Het eigendom van een patent ligt doorgaans op het niveau van de onderneming en is moeilijk door te vertalen naar bepaald (lokale) onderdelen of vestigingen van een onderneming. Ook universiteiten en instellingen vallen onder de eenheid onderneming. Niet alle patentaanvragen konden worden gekoppeld. Patentaanvragen die niet aan een onderneming uit het ABR gekoppeld konden worden, betreffen bijvoorbeeld patentaanvragen op persoonlijke titel. Deze aanvragen vallen onder 'onbekend' voor de bedrijfsclassificaties.

Procesverloop

De patentgegevens voor de jaartallen waarvan de status op StatLine nog niet definitief is worden opgehaald uit de PATSTAT online database. Additionele adresgegevens van Nederlandse aanvragers die ontbreken uit PATSTAT online worden geleverd door de Rijksdienst voor Ondernemend Nederland (RVO), welke enkel wordt gebruikt voor het koppelen van patentaanvragers met het bedrijvenregister maar niet wordt gepubliceerd. Het ontworpen algoritme koppelt waar mogelijk de aanvragen uit PATSTAT aan ondernemingen uit het ABR via naam-, adres- en woonplaatsgegevens van de aanvrager. Koppelingen waar

het algoritme geen duidelijke link kan leggen worden handmatig bekeken en beoordeeld. De resulterende dataset wordt bewerkt tot een microdataset welke wordt gecontroleerd op volledigheid en eventuele fouten. De dataset wordt ten slotte geaggregeerd op combinatie Standaard Bedrijfsindeling (SBI)/grootteklasse/land van zeggenschap (binnenland of buitenland) of op technologiegebied (WIPO - World Intellectual Property Organisation, of IPC-CBS - International Patent Classification CBS indikking).

Bijlagen

variabelenwaardenPAT.xlsx

3. Bestandsopbouw en toelichting

Bestandsopbouw

Onderstaand volgt een overzicht van alle variabelen.

VolgNr	Naam	Formaat
1	Auth	A2
2	Patstat-psn-id-crypt	F9
3	Og-id	F9
4	Docdb-family-id-crypt	F8
5	Status samw btl	F1
6	UCI	A10
7	SBI ABR	A6
8	GK SBS	A29
9	Technologiegebied IPC CBS	A2
10	Tech weight IPC CBS	F11
11	Technologiegebied WIPO	A2
12	Tech weight WIPO	F11

Toelichting op de variabelen

In deze paragraaf wordt informatie over de variabelen gegeven.

Auth

Indicator autoriteit aangevraagde patent

Definitie

De autoriteit (octrooibureau of instantie) verantwoordelijk voor het verwerken van een patentaanvraag.

Toelichting bij de definitie

EP staat voor European Patent Office (EPO). Een patent dat wordt toegekend door het EPO is geldig in elk van de landen die lid zijn van het EPO, mits de aanvrager zijn rechten in een land heeft gevalideerd. NL staat voor Octrooicentrum Nederland (OCNL). OCNL is de patentverlenende instantie voor Nederland en is onderdeel van de Rijksdienst voor Ondernemend Nederland (RVO). Toegekende patenten zijn alleen geldig in Nederland.

Toelichting bij het gebruik

Een patent kan aangevraagd worden bij zowel European Patent Office (EPO) als Octrooicentrum Nederland (OCNL). De patentaanvraag (vertegenwoordigd door docdb_family_id) komt dan voor onder beide classificaties.

Codelijst

Waarde	Label
EP	European Patent Office (EPO)
NL	Octrooicentrum Nederland (OCNL)

Patstat-psn-id_crypt

Identificatie van een patentaanvrager in de Patstat database

Definitie

Unieke identificatienummer van een persoonlijke aanvrager verbonden aan een patentaanvraag in de EPO Worldwide Patent Statistical Database (PATSTAT).

Toelichting bij de definitie

Elke gestandaardiseerde naam in de Patstat database (European Patent Office) krijgt een persoonlijke id toegewezen. Iedere aanvraag in de database is verbonden aan een PATSTAT Standardized Name (PSN)-id. De psn-id kan dus een niet gekoppelde onderneming zijn, maar ook een natuurlijk persoon.

Toelichting bij het gebruik

De psn_id is alleen weergegeven als er geen koppeling mogelijk was met het Algemeen Bedrijvenregister (ABR) van het CBS, en er dus geen onderneming verbonden is aan de patentaanvrager. Alle aanvragers vertegenwoordigd met een psn-id worden standaard gezet op SBI (Standaard Bedrijfsindeling 2008) = onbekend, grootteklasse = onbekend en land van zeggenschap = binnenland. Indien een koppeling met het ABR wel mogelijk was, wordt een lege waarde gegeven aan de psn_id. Het StatLine totaal van patentaanvragers kan worden berekend door alle unieke psn-id's op te tellen met alle unieke Og-id's (ondernemingsgroep).

Og-id

Identificatie van een ondernemingsgroep

Definitie

Het identificerende nummer van de meest omvattende verzameling van in Nederland gevestigde rechtspersonen waartussen overwegende zeggenschap bestaat.

Toelichting bij het gebruik

Alle patentaanvragers die gekoppeld kunnen worden aan het Algemeen Bedrijvenregister (ABR) hebben een bijbehorende Og-id (ondernemingsgroep). De SBI (Standaard Bedrijfsindeling 2008), grootteklasse en land van zeggenschap (binnenland of buitenland) worden bepaald door de kenmerken van de onderneming in het ABR en kunnen jaar-op-jaar veranderen. Indien een koppeling met het ABR niet mogelijk was, wordt een lege waarde gegeven aan de Og-id. Het StatLine totaal van patentaanvragers kan worden berekend door alle unieke psn-id's op te tellen met alle unieke Og-id's.

Docdb-family-id-crypt

Patentaanvraag bestaande uit een patentfamilie

Definitie

Unieke identificatienummer voor een verzameling aan documenten welke betrekking hebben op dezelfde uitvinding, oftewel een patentfamilie, wat gelijk is aan een patentaanvraag in de EPO Worldwide Patent Statistical Database (PATSTAT).

Toelichting bij het gebruik

Het StatLine totaal kan berekend worden door alle unieke Docdb-family-id's bij elkaar op te tellen. Elke id vertegenwoordigt een enkele aanvraag. Een aanvraag bevat dubbele rijen als er meerdere technologiegebieden aan dezelfde aanvraag verbonden zijn, welke worden gewogen. Daarnaast kunnen meerdere aanvragers betrokken zijn bij dezelfde aanvraag. Het optellen van de gewogen technologiegebieden geeft hetzelfde totaal als het optellen van de id's. Het jaar waarin de aanvraag voorkomt is het jaar waarin het patent voor het eerst is aangevraagd.

Status_samw_btl

Status samenwerking buitenland patentaanvraag

Definitie

Is er bij een patentaanvraag door een Nederlandse aanvrager samengewerkt met een buitenlandse partij?

Toelichting bij het gebruik

Indien de variabele gelijk is aan 1, is er bij de patentaanvraag door de Nederlandse aanvrager samengewerkt met het buitenland, anders is de indicator 0. De StatLine totalen voor patentaanvragen en -aanvragers waar is samengewerkt met het buitenland kan berekend worden door de unieke id's te tellen waarvan Status_samw_btl gelijk is aan 1. Voor aanvragen zijn dit de unieke Docdb-family-id's, voor aanvragers is dit een samentelling van unieke Patstat-psn-id's (ID nummer aanvrager in PATSTAT) en Og-id's (ondernemingsgroep).

UCI

Ultimate Controlling Institutional unit vestiging binnenland of buitenland

Definitie

Het land van uiteindelijke zeggenschap is het land waar de Ultimate Controlling Institutional Unit (UCI) is gevestigd.

Toelichting bij de definitie

Zeggenschap is mogelijkheid om het algemeen beleid van een onderneming te bepalen. De Ultimate Controlling Institutional Unit (UCI) is de institutionele eenheid, hoger in de keten van zeggenschap, die niet onder zeggenschap valt van een andere institutionele eenheid.

Toelichting bij het gebruik

Het land van zeggenschap is voor de patentenstatistiek onderverdeeld in uitsluitend binnenland of buitenland. Indien de aanvrager niet gekoppeld is aan het Algemeen Bedrijvenregister (ABR) en de aanvrager onbekend is, wordt de zeggenschap per definitie op Nederland gezet.

Codelijst

Waarde	Label
Binnenland	Binnenland
Buitenland	Buitenland
Onbekend	Onbekend

SBI_ABR

De economische activiteit volgens de Standaard Bedrijfsindeling 2008 zoals opgenomen in het ABR.

Definitie

De economische activiteit van een onderneming, volgens de Standaard Bedrijfsindeling 2008.

Toelichting bij de definitie

De Standaard Bedrijfsindeling (SBI) 2008 is de Nederlandse hiërarchische indeling van economische activiteiten die vanaf 2008 door het CBS wordt gebruikt om bedrijfseenheden in te delen naar hun hoofdactiviteit.

Toelichting bij het gebruik

Indien geen koppeling mogelijk is tussen de patentaanvrager en het Algemeen Bedrijvenregister (ABR) wordt de SBI op 'onbekend' gezet (code 999999).

GK_SBS

Grootteklasse aantal werkenden volgens Structural Business Statistics (SBS).

Definitie

Het aantal werkenden (werkzame personen al dan niet in loondienst) ingedeeld in een klasse (grootteklasse), volgens Structural Business Statistics (SBS).

Toelichting bij de definitie

De grootteklasse volgens het aantal werkenden, ook wel GK SBS (Structural Business Statistics) genoemd, is ingevoerd op grond van regels van de Europese Unie. Deze

grootteklasse-indeling wordt afgeleid van het aantal werkenden in een bedrijf, ongeacht of zij in loondienst zijn of anderszins betaalde arbeid verrichten (werkzame personen al dan niet in loondienst).

Toelichting bij het gebruik

Indien geen koppeling mogelijk is tussen de patentaanvrager en het Algemeen Bedrijvenregister (ABR) wordt de grootteklasse op 'Werkzame personen: onbekend' gezet.

Technologiegebied_IPC_CBS

CBS ingedikte versie van de International Patent Classification-code (IPC) voor technologiegebieden van patentaanvragen

Definitie

Classificatie die aangeeft op welk(e) technologiegebied(en) een patentaanvraag betrekking heeft, naar indeling van een door het CBS ingedikte versie van de International Patent Classification (IPC).

Toelichting bij de definitie

Door het CBS ingedikte versie van de International Patent Classification-code (IPC). De IPC is vastgesteld door de World Intellectual Property Organization. Het sectie-niveau bestaat uit de eerste drie karakters van de IPC-classificatie. International Patent Classification-codes worden volgens deze classificatie aan patentaanvragen toegekend vanuit technologisch perspectief door de patentverlenende organisatie. De code(s) die een patentaanvraag krijgt is afhankelijk van de functie van de uitvinding of het toepassingsgebied. De IPC is verdeeld naar secties, klassen, subklassen, groepen en subgroepen (in totaal beslaat de IPC zo'n 70 000 categorieën). Op basis van de IPC heeft het CBS zelf aggregaten samengesteld.

Toelichting bij het gebruik

De codes vallen in de volgende categorieën: A Menselijke Levensbehoeften - B Bewerkingen; transport - C Chemie; metallurgie - D Textiel; Papier - E Vaste constructies - F Werktuigbouwkunde; verlichting - G Meet- en regeltechniek - H Elektriciteit.

Tech_weight_IPC_CBS

Gewogen IPC-CBS code toegekend aan patentaanvraag (International Patent Classification)

Definitie

Het gewicht dat aan een International Patent Classification (IPC) classificatie naar CBS indikking kan worden toegekend voor een patentaanvraag.

Toelichting bij de definitie

Elke patentaanvraag (Docdb-family-id) bevat als patentfamilie meerdere onderliggende IPC-codes (International Patent Classification). Gebaseerd op hoe vaak een IPC-code voorkomt in een patentaanvraag wordt een gewicht toegekend aan de betreffende IPC-CBS indikking.

Toelichting bij het gebruik

Voor elke unieke aanvraag (Docdb-family-id) tellen de onderliggende gewichten altijd op tot 1. Ter illustratie kunnen er voor een aanvraag 3 codes voorkomen van het ene type en 1 code

van het andere type: het verdeelde gewicht is dan 0,75-0,25. Door alle fracties bij elkaar op te tellen kan het StatLine totaal van de patentaanvragen worden berekend.

Technologiegebied_WIPO

Classificatie naar indeling WIPO (World Intellectual Property Organisation) voor technologiegebieden van patentaanvragen

Definitie

Classificatie die aangeeft op welk(e) technologiegebied(en) een patentaanvraag betrekking heeft, naar indeling van de technologie-concordantietabel van de World Intellectual Property Organisation (WIPO), welke de symbolen van de Internationale Patenten Classificatie (IPC) verbindt met vijfendertig technologiegebieden.

Toelichting bij de definitie

De WIPO-concordantietabel (World Intellectual Property Organisation) wordt regelmatig bijgewerkt aan de hand van herzieningen van de IPC (International Patent Classification). Iedere patentaanvraag wordt aan de hand van de IPC-codes waarop de aanvraag betrekking heeft toegekend aan één of meerdere hoofdgebieden. De hoofdgebieden zijn Elektrotechniek, Instrumenten, Chemie, Werktuigbouw en Overige technologiedomeinen.

Toelichting bij het gebruik

De codes vallen in de volgende categorieën: I Elektrotechniek - II Instrumenten - III Chemie - IV Werktuigbouw - V Overige technologiedomeinen.

Tech_weight_WIPO

Gewogen WIPO code toegekend aan patentaanvraag (World Intellectual Property Organisation)

Definitie

Het gewicht dat aan een World Intellectual Property Organisation (WIPO) classificatie kan worden toegekend voor een patentaanvraag.

Toelichting bij de definitie

Het gewicht dat aan de betreffende WIPO code (World Intellectual Property Organisation) kan worden toegekend voor een patentaanvraag. Elke patentaanvraag (Docdb_family_id) bevat als patentfamilie meerdere onderliggende WIPO-codes (World Intellectual Property Organisation). Gebaseerd op hoe vaak een WIPO-code voorkomt in een patentaanvraag wordt een gewicht toegekend aan de betreffende WIPO classificatie.

Toelichting bij het gebruik

Voor elke unieke aanvraag (Docdb_family_id) tellen de onderliggende gewichten altijd op tot 1. Ter illustratie kunnen er voor een aanvraag 3 codes voorkomen van het ene type en 1 code van het andere type: het verdeelde gewicht is dan 0,75-0,25. Door alle fracties bij elkaar op te tellen kan het StatLine totaal van de patentaanvragen worden berekend.

Versiegeschiedenis

<i>Bestandsnaam</i>	<i>Reden</i>
PAT2013V1 t/m PAT2018V1	Definitief
PAT2019V1	Nader voorlopig
PAT2020V1	Voorlopig
PAT2019V2	Herziene versie, vervangt 2019V1
PAT2020V2	Herziene versie, vervangt 2020V1