

TalentAnalytics

Eerste pseudonymisatie

Versiegeschiedenis			
0.1	07/06//2024	VST	Eerste versie
0.2	07/06/2024	VST + NDE	Verwerking opmerkingen van Nicolas de Timmerman (DPO talent.brussels)
0.3	31/10/2025	MXM	Aanmaak Nederlandse vertaling en verduidelijking in te vullen gegevens.

1 Inleiding

In het kader van de verzameling van talentAnalytics-gegevens moeten de partners die hun gegevens naar talent.brussels sturen via de gewestelijke dienstenintegrator Paradigm, de gegevens pseudonimiseren alvorens ze naar Paradigm te verzenden, om de risico's op heridentificatie van personen te beperken.

Dit document beschrijft de te implementeren pseudonimiseringsmethode om zowel aan de behoeften van TalentAnalytics als aan de vereisten inzake gegevensbescherming te voldoen.

2 Methode

De te gebruiken methode bestaat erin een tekenreeks van 21 karakters te creëren en deze tekenreeks te hashen met behulp van een hashing-algoritme dat in punt 2.2 wordt gespecificeerd.

2.1 Tekenreeks

De gegevens die in het hashing-algoritme moeten worden gebruikt, bestaan uit 3 delen:

- De eerste 5 karakters van de voornaam
- De 11 cijfers van het rijksregisternummer
- De eerste 5 karakters van de naam

2.1.1 Naam en voornaam

De eerste 5 karakters van de naam en voornaam worden gebruikt.

- Karakters met accenten, cedille, ... worden vervangen door het karakter zonder accent
- Koppeltekens en spaties worden verwijderd

De eerste 5 karakters van de officiële naam en voornaam worden gebruikt, waarbij de volgende wijzigingen worden toegepast:

- Andere karakters dan de 26 basisletters van het Latijnse alfabet worden verwijderd: koppeltekens, spaties, apostroffen, enz.
- Ligaturen worden gescheiden: Æ => AE; Œ => OE

- Karakters met accenten, een cedille of enig ander diakritisch teken¹ worden vervangen door hun basisversie: ç => c; É => E; ï => i; Ô => O

De karakters worden omgezet in hoofdletters. Als de resulterende tekenreeks minder dan 5 karakters lang is, worden er zoveel X'en aan het einde toegevoegd tot 5 karakters worden bekomen.

Voorbeelden:

- Luc ➔ LUCXX
- De Brouckère ➔ DEBRO
- Céline ➔ CELIN

2.1.2 Rijksregisternummer

Alleen de cijfers waaruit het bestaat, worden gebruikt.

Voorbeeld: 42.01.22.051-81 ➔ 42012205181

2.2 Te gebruiken algoritme

Op aanbeveling van het FIDUS-team van Paradigm is het te gebruiken algoritme voor het hashen van het berekende veld **SHA3-256**.

De definitie van dit algoritme is te vinden op dit adres: [SHA-3 — Wikipedia \(wikipedia.org\)](https://en.wikipedia.org/wiki/SHA-3)

Een online implementatie is hier te vinden: [SHA3-256 - Online Tools \(emn178.github.io\)](https://emn178.github.io/online-tools/sha3-256/)

3 Voorbeelden

Brongegevens:

Luc - Delbrouck - 42.01.22.051-81

Te coderen in kolom unique_agent_id:

LUCXX41012205181DELBR

Resultaat encoding door FIDUS:

f378edf263d60a8265ac9a046a81dobo828f4ce2a8a68e6591b55b757aeef24

Brongegevens:

Vincent - Streignard - 73.10.22.051-81

Te coderen in kolom unique_agent_id:

VINCE73102205181STREI

Resultaat encoding door FIDUS:

f0963a71f09577f558d91812cc8cfodeeca5a4e3b23b59a9a3ef4b2dodoa7cf4

¹ [Diakritisch teken - Wikipedia](https://en.wikipedia.org/wiki/Diacritic)

Het zo verkregen pseudoniem hoeft niet noodzakelijkerwijs vóór elke verzending naar Paradigm voor TalentAnalytics te worden berekend. Het kan immers vóór de eerste verzending worden berekend en veilig worden opgeslagen in het informaticasysteem voor personeelsbeheer van de partner.

Het is echter belangrijk dat het niet voor andere doeleinden wordt gebruikt dan de overdracht van gegevens aan TalentAnalytics, en dat het wordt megedeeld aan de dienstenintegrator Paradigm en niet aan talent.brussels.

Paradigm koppelt immers aan dit pseudoniem een ander pseudoniem, dat het volledig willekeurig toekent, en stuurt alleen dit laatste door naar talent.brussels.

De werkelijke identiteit van de personen mag en kan niet worden megedeeld aan Paradigm of aan talent.brussels in het kader van TalentAnalytics.

