Kennisnet

Nummervoorziening voorschriften

Pseudonimisering in de leermiddelenketen

Laat ict werken **voor het onderwijs**

Inhoud

Document informatie

Status

Auteur	Kennisnet, Marc Fleischeuers
Versie	1.0.10
Versiedatum	15 maart 2017
Status	Goedgekeurd

Versiehistorie

Versie	Datum	Auteur	Beschrijving
0.9	6 april 2016	Marc Fleischeuers	Samengesteld uit eerdere documenten
0.91	18 april 2016	Marc Fleischeuers	Naar aanleiding van Edu-K KAT overleg
0.92	11 mei 2016	Marc Fleischeuers	Encryptieeisen VS, Voorschriften 5 en 6 verder toegelicht, voorschrift 1 uitgewerkt
1.0	12 juli 2016	Marc Fleischeuers	
1.0.1	21 december 2016	Marc Fleischeuers	Commentaar Voorschrijft 1, Voorschrift 6 verwerkt
1.0.9	23 januari 2017	Marc Fleischeuers	Wijziging wetgeving (introductie stampseudoniem), voorschrift 9
1.0.10	15 maart 2017	Marc Fleischeuers	Nav KAT 8-3-17: Voorschrift 1, 5, 9 aanscherping; voorschrift 6 herschreven.
1.0.11	26 oktober 2017	Marc Fleischeuers	ECK-ID specifieke eisen afgesplitst

Distributie en goedkeuring

Versie	Datum	Goedgekeurd (indien van toepassing)
1.0.10	21-3-2017	Tactisch overleg Toegang tot leermateriaal

1. Voorschriften voor het koppelen met de Nummervoorziening

In dit hoofdstuk worden de maatregelen rondom de beveiliging beschreven van het ketenpseudoniem dat gebruikt zal worden in de educatieve leermiddelenketen (ECK ID). In de Algemene Verordening Gegevensbescherming (AVG[1]) wordt pseudonimiseren van persoonsgegevens voorgesteld als privacybeschermende maatregel. Er worden voorwaarden gesteld aan de pseudonimisering, zoals te zien in de definitie:

> *„pseudonimisering": het verwerken van persoonsgegevens op zodanige wijze dat de persoonsgegevens niet meer aan een specifieke betrokkene kunnen worden gekoppeld zonder dat er aanvullende gegevens worden gebruikt, mits deze aanvullende gegevens apart worden bewaard en technische en organisatorische maatregelen worden genomen om ervoor te zorgen dat de persoonsgegevens niet aan een geïdentificeerde of identificeerbare natuurlijke persoon worden gekoppeld; (AVG, Art 4 lid 5)*

Pseudonimisering is hier een methode die er voor zorgt dat de persoon niet identificeerbaar is. Over identificeerbaarheid zegt de verordening het volgende:

> *Om te bepalen of een natuurlijke persoon identificeerbaar is, moet rekening worden gehouden met alle middelen waarvan redelijkerwijs valt te verwachten dat zij worden gebruikt door de verwerkingsverantwoordelijke of door een andere persoon om de natuurlijke persoon direct of indirect te identificeren, bijvoorbeeld selectietechnieken. Om uit te maken of van middelen redelijkerwijs valt te verwachten dat zij zullen worden gebruikt om de natuurlijke persoon te identificeren, moet rekening worden gehouden met alle objectieve factoren, zoals de kosten van en de tijd benodigd voor identificatie, met inachtneming van de beschikbare technologie op het tijdstip van verwerking en de technologische ontwikkelingen. (AVG overweging 26)*

De hieronder genoemde maatregelen zijn de technische en organisatorische maatregelen om ervoor te zorgen dat de gepseudonimiseerde gegevens van leerlingen en docenten[2] niet aan de persoon zelf kunnen worden gekoppeld, tenzij hiervoor onredelijk zware middelen ingezet worden.

Met deze beveiligingsmaatregelen worden de volgende doelen gerealiseerd:

1. ECK IDs voor leerlingen van het PO, VO, MBO zijn blijvend uniek binnen het onderwijs.
2. ECK IDs voor leerlingen van het PO, VO, MBO zijn persistent gedurende tenminste het verblijf van de leerling in een onderwijssector.
3. ECK IDs zijn niet direct herleidbaar tot de persoon en geven geen informatie over de persoon.
4. Maatregelen zijn gericht op voorkomen van onbevoegd gebruik en voorkomen van misbruik van het ECK ID
5. De gebruikte encryptie- en hashing-algoritmen worden breed toegepast in productiesystemen, zijn goed onderzocht en blijven voor de voorziene toekomst, 5 à 10 jaar, nog actueel[3]. Zie hiervoor met name voorschriften 2, 4, 5, 6 en 9 op pagina 4 en verder.

De maatregelen worden beschreven aan de hand van de verwerkingen die plaats vinden tijdens het aanmaken en gebruik van het ECK ID. Een en ander is geïllustreerd aan de hand van de basis use case "Genereer een ECK ID" in het document "Principes en processen".

De hier geformuleerde maatregelen omtrent encryptie en hashing zijn gebaseerd op "ICT richtlijnen voor Transport Layer Security", Nationaal Cyber Security Centrum van het Ministerie van Justitie, "Gebruik en Achtergrond Digikoppeling Certificaten", Logius, versie 1.2.1 (definitief), en "Algorithms, key size and parameters report - 2014", European Union Agency for Network and Information Security (ENISA), november 2014. De

[1] Verordening (EU) 2016/679 van het Europees Parlement en de Raad van 27 april 2016.

[2] Leerlingen en docenten zijn hier bedoeld als alle onderwijsvolgers en medewerkers van onderwijsinstellingen.

[3] "Algorithms, key size and parameters report – 2014", European Union Agency for Network and Information Security (ENISA), 2014

aanbevelingen over sleutellengten en algoritmen uit het ENISA rapport komen overeen met diverse andere aanbevelingen van Europese en Amerikaanse[4] agentschappen.

1.1. *Voorschrift NV-1*: 1e niveau hashing van de basis van het stampseudoniem

Hiervoor wordt gebruik gemaakt van scrypt. Deze functie heeft een aantal parameters waarmee de zwaarte van het berekenen van de hash wordt gestuurd. De zwaarte is zodanig gekozen dat het normale werk, het aanvragen van een ECK ID bij de aanmelding van een leerling, niet wordt gehinderd terwijl het misbruik (aanleg van een vertaaltabel) wordt ontmoedigd.

Op basis van overwegingen dat de 1e niveau hashing server-side wordt toegepast (dwz in een omgeving waar veel geheugen aanwezig is) worden deze parameters gebruikt:

Parameter	Waarde	Toelichting
N	2^{17} = 131072	Bepaalt wat het geheugenbeslag van de individuele operaties gaat zijn.
r	8	Factor voor geheugenbeslag.
p	4	Parameter voor parallelle verwerking en dus voor de tijd die de berekening gaat kosten.
salt	string	De te gebruiken salt wordt beschikbaar gemaakt als onderdeel van het aansluitproces.

Figuur 1 Voorgeschreven parameters voor scrypt

Met deze parameters zal het geheugenbeslag voor de aanvraag van een stampseudoniem ongeveer 130MB en duurt op een core i5 (single core) ongeveer 10 seconden. In het originele paper van scrypt (http://www.tarsnap.com/scrypt/scrypt.pdf) wordt geadviseerd om N = 2^{14} of N = 2^{16} te kiezen. Voor normaal gebruik op client systemen (PC, laptop) zijn dit nog steeds veel gebruikte waarden. Voor de server-omgeving waarin deze functie wordt gebruikt geldt dat deze meer geheugen en meer CPU cores beschikbaar heeft, vandaar de iets zwaardere parameters.

Nota bene: er zijn scrypt bibliotheken die de gekozen parameters opnemen in de output, bijvoorbeeld de veel gebruikte Java SCrypt library[5] die als standaard uitvoer heeft “$s0$params$salt$hashwaarde” . Voor verwerking in de Nummervoorziening dient *alleen* de hashwaarde te worden gebruikt, uitdrukkelijk zonder deze parameters.

De uitvoer van de scrypt functie, de hashwaarde, is een binaire bytearray. Dit getal wordt voor onderlinge uitwisseling geconverteerd naar een 64 karakters lange hexadecimale string.

1.2. *Voorschrift NV-2*: Toepassen van Edukoppeling.

Edukoppeling wordt toegepast voor server-naar-server communicatie tussen leerlingadministratie client en Nummervoorziening. Hiervoor geldt:

a. Er wordt gebruik gemaakt van Edukoppeling 1.2 of een latere, door de Standaardisatieraad vastgestelde versie.

Een nadere beschrijving van de eisen die in Edukoppeling 1.2 worden gesteld zijn:

b. Het gebruikte certificaat van de service client bevat het OIN van de leerlingadministratie-leverancier en is verkregen bij een Digikoppeling CSP.

c. Partijen die dienst afnemen hebben de plicht om hun certificaat in te laten trekken indien dit niet meer gebruikt wordt of gecompromitteerd is.

[4] Het Information Assurance Directorate van het NSA heeft voor een specifieke, beperkte groep commercieel verkrijgbare pakketsoftware met encryptiefunctionaliteit bepaald dat deze alleen met een vergunning geëxporteerd mag worden. Voor de hier besproken algoritmes geldt in het algemeen dat ze niet onder deze bepaling vallen. Zie http://www.bis.doc.gov/index.php/regulations/export-administration-regulations-ear en http://www.bis.doc.gov/index.php/forms-documents/doc_download/951-ccl5-pt2

[5] https://github.com/wg/scrypt

d. De WS-Addressing headers van het bericht bevatten OINs van onderwijsinstelling en de Nummervoorziening. Het OIN van de onderwijsinstelling[6] is een geldig OIN volgens de definitie van Digikoppeling.

e. Het bericht is 2W-BE, en niet ondertekend en niet encrypted.

1.3. *Voorschrift NV-3*: De hashing van de Nummervoorziening.

Hiervoor wordt gebruik gemaakt van SHA-512 hash algoritme. Dit algoritme maakt een hash van 512 bytes output. Deze waarde wordt gezien als toekomstvast en wordt aangeraden in het rapport “Algorithms, key size and parameters report – 2014”, European Union Agency for Network and Information Security (ENISA), 2014. In het Afsprakenstelsel eHerkenning wordt gesteld dat een pseudoniem een hexadecimale string van 32 karakters is, dat wil zeggen 128 bits. Dit is een tamelijk korte sleutel en niet heel toekomstvast. Om deze reden wordt gekozen om een langere sleutel (512 bits, 128 hexadecimaal karakters) te gebruiken.

Voor communicatie converteert de nummervoorziening de hashwaarde naar een 128 karakters lange hexadecimale letterreeks.

De gebruikte salt-waarden wordt veilig en duurzaam opgeslagen met de minimale set aan toegangsrechten: alleen leestoegang voor het serviceaccount dat gebruikt wordt om de hashingoperatie uit te voeren.

1.4. *Voorschrift NV-4*: Toepassen van TLS tussen LAS en Nummervoorziening

Dit voorschrift wordt toegepast voor systeem – systeem verbindingen zoals tussen LAS en de Nummervoorziening. Voor deze verbindingen geldt dat er voorschriften aan zowel de client (LAS) als de server kunnen worden opgelegd. Hiervoor wordt het gebruiksadvies van de ICT-beveiligingsvoorschriften van het NCSC gevolgd voor het scenario waarin controle is over client en server. Dit betreft dus maatregelen die de Nummervoorziening zelf implementeert en maatregelen die ketenpartijen implementeren in hun client.

- Gebruik alleen TLS versie 1.2 of hoger.
- Kies GOEDE cyphersuites voor de server. Ondersteun niet meer cyphersuites dan alleen deze GOEDE suites. Zie de ICT-beveiligingsvoorschriften van het NCSC voor een uitputtende lijst.
- Kies voldoende lengte van parameters en sleutels. Zie de ICT-beveiligingsvoorschriften van het NCSC voor aanbevelingen.

[6] Het OIN is in Digikoppeling gedefinieerd als numeriek veld. Een BRIN bevat twee letters, daarom is het in de educatieve sector gebruikelijk om het OIN als alfanumeriek te gebruiken.