Computest
always on.

Whitepaper **Security**

Deze whitepaper is tot stand gekomen in samenwerking met: **www.vka.nl**

VERDONCK KLOOSTER & ASSOCIATES

Privacy-vriendelijk testen? Geen punt!

INTRO

Testen met persoonsgegevens, mag dit nu wel of mag dit niet? Uiteindelijk is het antwoord: als het noodzakelijk is wel, maar eigenlijk wil je dit niet en is het ook zelden nodig. Hoe dan ook: de tijd van ‘een kopietje van productie’ is voorbij!

Testen met persoonsgegevens, mag dit nu wel of mag dit niet? Uiteindelijk is het antwoord: als het noodzakelijk is wel, maar eigenlijk wil je dit niet en is het ook zelden nodig. Hoe dan ook: de tijd van 'een kopietje van productie' is voorbij!

In dit artikel bundelen Computest en Verdonck, Klooster & Associates (hierna: VKA) hun kennis en ervaring om de werelden van beveiliging, privacy en testmanagement bij elkaar te brengen. Computest als specialist op het gebied van security en testautomatisering, VKA als privacy-expert. We beschrijven enerzijds wat de beperkingen en mogelijkheden zijn vanuit de wet- en regelgeving op het gebied van beveiliging en privacy. Anderzijds beschrijven we hoe professioneel testmanagement er anno nu uitziet en hoe dit past bij privacy-vriendelijk testen. We sluiten af met een aantal praktische uitgangspunten.

Mag het wel?

Allereerst de vraag: mag je testen met persoonsgegevens? De privacy-toezichthouder (Autoriteit Persoonsgegevens) stelt: "Nee, dat mag niet. Een organisatie die uw persoonsgegevens heeft, mag deze gegevens niet gebruiken voor een ander doel. Het testen van informatiesystemen mag alleen met fictieve (verzonnen) gegevens of met persoonsgegevens die weinig risico met zich meebrengen. Dat zijn bijvoorbeeld openbare gegevens, zoals van publieke websites."

Nu staat nergens in de privacywetgeving letterlijk dat het verboden is om te testen met productiegegevens. Wel zegt de wet dat je gegevens alleen mag gebruiken voor het doel waarvoor je ze hebt verkregen, en voor doelen die in duidelijk direct verband daarmee staan (art. 8 en 9 Wet bescherming persoonsgegevens, de Wbp). Nu heeft het testen van een omgeving waarbinnen je persoonsgegevens wilt gaan gebruiken wel degelijk enig verband met het daadwerkelijk (productie) gebruik van die omgeving. Dus als er toestemming is voor het productiegebruik waarom dan ook niet voor validatie en acceptatie van de software die dat gaat doen?

> Tegelijkertijd is de gedachte achter de privacywetgeving dat de hoeveelheid persoonsgegevens die worden gebruikt om te testen moeten worden ingeperkt tot een onvermijdelijke proportie. Maar waar in een ontwikkel- en testtraject gaat het vermijdelijke gebruik over in het onvermijdelijke gebruik?

Het meest privacy-vriendelijke uitgangspunt is daarmee dat je daar waar het gebruik van productiedata niet nodig is (bijvoorbeeld in een ontwikkel- en testtraject) nepdata moet gebruiken. En alleen bij die tests waar de echte data gebruikt móet worden, bijvoorbeeld om de kwaliteit van een bestaande koppeling ten behoeve van een migratie vast te stellen, kan het gebruik van 'echte gegevens' noodzakelijk zijn.

Wil je wel?

Dan de vraag: wil je wel testen met productiegegevens? Dit is een vraag niet te maken heeft met privacy, maar veel meer met de vraag hoe je kwalitatief goed kunt testen. Daarom is het belangrijk eerst te kijken: waar heb je testdata voor nodig? Veelal zal het zijn om goedsituaties, de zogenaamde happy flow te testen.

Daarnaast wil je foutsituaties en bijbehorende foutmeldingen testen en soms ook grenswaarden testen. Hiervoor is geen productiedata nodig en zeker als er testautomatisering gaat plaatsvinden is het vanwegede voorspelbaarheid en onderhoudbaarheid van je geautomatiseerde testscript juist verstandig om gebruik te maken van eigengemaakte en controleerbare testdata.

Deze testdata wordt elke testcyclus zelf aangemaakt door middel van gebruik van de applicatie onder test. Dit heeft als additioneel voordeel dat er soms fouten in de applicatie al onbedoeld worden opgespoord. Hoewel het aanmaken van testdata geldt als best practice, kiezen veel mensen nog steeds voor de op het eerste oog 'makkelijkere en snellere' variant: productiedata.

Niet makkelijk en sneller

Dat 'makkelijk en sneller' is echter relatief: het selecteren van de juiste testcases uit een grote hoeveelheid productiedata is een stuk lastiger dan bij een beperkte set testdata. Hier zorgt het gebruik van productiedata regelmatig voor tijdsverlies. Daarnaast zie je vaak dat een bepaalde case niet in de productiedata zit omdat:

1. Die situatie zich (nog) niet voorgedaan heeft, of;
2. Het gaat over nieuwe functionaliteit die wordt ingebouwd, waarvoor nog geen testdata beschikbaar kan zijn.

In dat geval dan zal er toch nog nieuwe data moeten worden gemaakt, of bestaande data worden gemodificeerd.

Het 'makkelijk en sneller' geldt ook niet als men voor het testen gebruik wil maken van testautomatisering. De kwaliteit van de testdata in een productiebestand is eenvoudigweg minder goed dan de kwaliteit van gegenereerde testdata, omdat het onbekend is welke van de flows door de applicatie wel en niet geraakt zullen worden door productiedata. Gegenereerde testdata zorgen dat de geautomatiseerde testscripts een stuk robuuster zijn. Daarnaast is het ook eenvoudiger om gevonden fouten tijdens de geautomatiseerde testuitvoer te analyseren indien er gebruik wordt gemaakt van eigengemaakte testdata, dit omdat er minder tijd verloren gaat met het beoordelen van de gebruikte testdata. Zit de fout in de applicatie of heb ik verkeerde data gebruikt?

Vaak wordt gezegd dat het gebruik van productiedata noodzakelijk is ten bate van bijvoorbeeld load- en stresstesten, dit omdat voor dit type testen een groot volume data nodig is. Het handmatig aanmaken van deze data wordt gezien als tijdrovend en daarom ondoenlijk. Naar onze mening is het lang niet altijd noodzakelijk om terug te vallen op productiedata. Er zijn verschillende oplossingen te bedenken om voldoende testdata te creëren. Door slim gebruik te maken van testautomatiseringstools en de mogelijkheid om datadriven te testen, kunnen geautomatiseerd grote hoeveelheden testdata gecreëerd worden. Mocht dit geen optie zijn, dan zij er legio tools en websites beschikbaar om in veel verschillende formaten testdatasets te genereren die geïmporteerd kunnen worden, waardoor toch een relevante hoeveelheid testdata ontstaat. Pas als men zich moeite heeft getroost om deze alternatieven te onderzoeken, dan blijft het anonimiseren van productiedata over.

Beveiliging tijdens het testen

Laten we ervan uitgaan dat in het belangrijkste deel van de ontwikkel- en testcyclus 'echte' persoonsgegevens niet nodig zijn. Dan blijft er altijd nog een deel van de testwerkzaamheden over waarvoor het wel 'onvermijdelijk' is dat met persoonsgegevens wordt gewerkt. Daar ontstaan nog regelmatig datalekken omdat de gegevens niet goed genoeg zijn beveiligd.

In de gevallen dat het gebruik van echte persoonsgegevens bij het testen noodzakelijk is, moeten deze gegevens 'adequaat worden beschermd', zoals de Wet Bescherming Persoonsgegevens dit stelt.

Wanneer je in je testomgeving productiegegevens opneemt, dan moet de testomgeving aan beveiligingseisen voldoen die vergelijkbaar zijn met die van je productieomgeving. In de praktijk zijn dan de volgende beveiligingsmaatregelen van belang:

- Beperken van de toegang – wie kunnen en mogen erbij? Bij het ontwikkel- en testproces zijn vaak ook externe partijen betrokken. Deze ontwikkelaars zitten remote of on-site, wat specifieke eisen stelt aan de toegangsbeveiliging, op het niveau van fysieke toegang, netwerk, database of applicatie.
- Verwijderen van (kopieën van) testdata – verwijder de (kopieën van) testdata na gebruik, ook de backups. Verplicht ook de externe ontwikkelaars hiertoe en laat hen bewijzen dat alles is vernietigd.
- Maak de testgegevens niet online toegankelijk en als dit al nodig is, maak geen stomme configuratiefouten: scherm de toegang dan goed af en schakel bijvoorbeeld ook de indexering voor zoekmachines in Google uit.
- Het is tevens verstandig om de logging te activeren op het gebruik van de testdata. Zo kan je zien of de gevoelige testgegevens 'per ongeluk' zijn gebruikt voor verkeerde doeleinden en ben je in staat, in geval van incidenten, de impact beter in te schatten zodat je beter het incident kunt verhelpen.

En de belangrijkste: security awareness! Als de testers en ontwikkelaars niet begrijpen wat het belang is van de beveiliging van persoonsgegevens en wat de consequenties zijn van een datalek, dan is het gebruik van productiedata buiten de productieomgeving een risicovolle onderneming.

Samengevat

Testen met gegenereerde in plaats van 'echte' persoonsgegevens is veiliger en draagt bij aan een hogere kwaliteit van testen. En dus uiteindelijk van de software. Testdata kan zo worden gegenereerd dat deze alle flows in de applicatie zal triggeren, waar productiedata vaak bepaalde cases mist omdat die nog nooit voorgekomen zijn. Daarnaast brengt het verschepen van productiedata naar testomgevingen kopzorgen met zich mee op het gebied van beveiliging.

Hoe kan je veilig en privacy-vriendelijk testen:

- Stel vast voor welke tests het gebruik van productiegegevens met persoonsinformatie vermijdelijk of onvermijdelijk is. Neem hierbij niet te snel aan dat je productiegegevens nodig hebt.
- Gebruik zo min mogelijk productiegegevens en gebruik zo veel mogelijk representatieve, gegenereerde testdata.
- Daar waar gebruik productiegegevens onvermijdelijk is: beperk de toegang, zorg voor traceerbaarheid en verwijder zo snel mogelijk en anonimiseer waar dit kan. Maak voor dit laatste gebruik van technische pseudonimiserings- en anonimiseringstools.

Bronnen

- http://www.binnenlandsbestuur.nl/digitaal/kennispartners/info-support/software-testen-met-persoonsgegevens-mag-dat-nog.9544928.lynkx
- http://blog.iusmentis.com/2016/03/31/mag-systemen-testen-productiedatapersoonsgegevens/
- https://autoriteitpersoonsgegevens.nl/nl/over-privacy/persoonsgegevens/beveiliging-van-persoonsgegevens
- https://www.enterprisetimes.co.uk/2016/11/14/michael-page-data-leakblamed-capgemini/



Computest
always on.

info@computest.nl
+31(0)88 733 13 37
www.computest.nl

We zijn een team van gepassioneerde en ervaren technisch specialisten die applicaties en infrastructuren optimaal laten werken. Wij geloven in geïntegreerde quality assurance en bieden daarom diensten op het gebied van performance, security en functionele testautomatisering.

In alles wat we doen worden we gedreven door een grenzeloze passie voor kwaliteit. Daarom werken we voor iedere sector samen in kleine, gespecialiseerde agile teams. Daarmee houden we de lijnen kort zodat we de beste resultaten behalen.