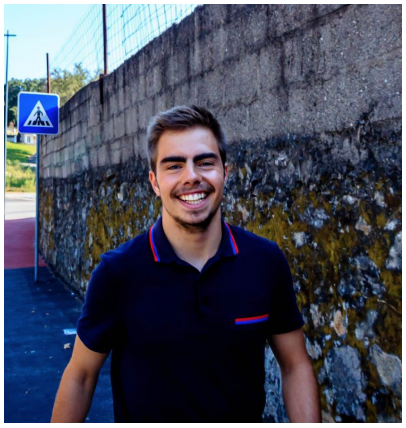
Escola de Engenharia
Universidade do Minho

DEPARTAMENTO DE ENGENHARIA INFORMÁTICA
Mestrado em Engenharia Informática
Engenharia de Segurança

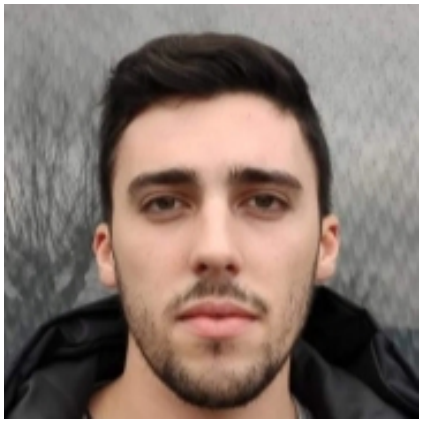
Aula 6

16 de Março de 2020

Grupo 1



Ricardo Pereira a73577



Tiago Ramires pg41101

Braga, 23 de Março de 2020

1. RGPD (Regulamento Geral de Proteção de Dados)

Pergunta P1.1

A pseudonimização é uma técnica que tem como propósito permitir a ocultação da identidade de um indíviduo num qualquer sistema de dados. Essencialmente, a informação que nos é apresentada hoje em dia, na grande maioria das vezes é guardada sem qualquer proteção e sempre que alguém acede aos dados consegue fazê-lo percebendo-os por inteiro e mais importante, consegue perceber a quem é que os dados se referem e pertencem. Por esses e por outros motivos, o RGPD apareceu para tentar regulamentar esta área das tecnologias de informação de forma a que as organizações que lidam constantemente com dados, principalmente dados de cidadãos comuns, tenham regras para cumprir, impossibilitando que as mesmas os utilizem de forma errada. Ora, como esperado, o RGPD trouxe bastantes entraves a estas organizações e a pseudononimização permite-lhes "fugir"a algumas obrigações a que estão sujeitas.

Os dois principais objetivos da pseudononimização são:

- um agente externo (algo que não seja o controlador ou o processador) não deve conseguir identificar facilmente o pseudónimo nem sequer relacionar o contexto dos dados em questão;
- um agente externo não deve conseguir reproduzir facilmente um pseudónimo.

Obviamente, para fazer a ligação dos pseudónimos aos utilizadores é necessário que o sistema conheça um mecanismo para fazer a associação, sendo esse mecanismo conhecido pelo controlador/processador de dados.

Para esta técnica de pseudononimização é possível derivar pseudónimos a partir de funções de *hash* contudo, existem alguns contras e alguns aspetos menos bons a ter em conta que podem eventualmente ser melhorados. Por exemplo, com a geração de pseudónimos usando *hashing* sem chave, a primeira propriedade falada anteriormente não se verifica visto que é possível verificar a correspondência de um pseudónimo relativamente a um identificador. A segunda propriedade também falha pois se aplicarmos uma mesma função de hash ao mesmo identificador, obtém-se o mesmo pseudónimo. Não obstante, conseguem resolver-se alguns destes problemas com o uso de *hashing* com *salts* ou chaves - a primeira propriedade valida-se pois sem a chave, não é possível associar o pseudónimo ao identificador e a segunda também se valida pois com o mesmo identificador podem gerar-se vários pseudónimos.

Outra forma de se fazer pseudononimização é através de cifras, simétricas ou assimétricas. sendo que nas simétricas é criado um criptograma que será usado como pseudónimo cujas partes que querem aceder aos dados, devem possuir a chave de desencriptação que obviamente

não pode ser conhecida por terceiros. Nas assimétricas, podem ser utilizados os métodos tradicionais de encriptação com par chave pública/privada, com a agravante de que estas chaves necessitam de ser grandes para que haja maior segurança.

Este conceito está em constante evolução e como podemos ver no artigo que se leu, as organizações que precisem de o implementar podem até criar a sua própria solução. Ainda assim, existem outras soluções tais como *tokenisation*, *masking*, *scrambling*, *blurring*, entre outras. *Tokenization* é um método que já é amplamente utilizado nas instituições bancárias onde o pseudónimo é substituido por um *token* que é gerado aleatoriamente, não havendo assim qualquer relacionamento com o identificador ou com os dados para os quais foram gerados.

Pergunta P1.2

O assunto que se aborda neste caso de uso é a gestão das folhas com informação financeira de empregados de algumas organizações. De acordo com o artigo em questão, o processamento de folhas de salários, benefícios e assuntos de segurança social obriga a que as empresas obedeçam e cumpram regras de confidencialidade, contudo não há nada regulamentado que indique os procedimentos a ter quanto à retenção e destruição desses dados.

Para ser feita uma melhor análise ao problema, importa diferenciar as várias operações de processamento:

- **processamento de dados pessoais** - informações de contacto, dados de segurança social, informação salarial, etc;
- **finalidade** - pagamento de salário, submissão de documentos para segurança social, etc;
- **assunto dos dados** - empregados da organização;
- **meios de processamento** - sistema de recursos humanos da organização;
- **destino dos dados** - várias entidades governamentais como segurança social, finanças, etc;
- **processador de dados utilizado** - nenhum.

Com a especificação anterior está-se a limitar e ao mesmo tempo a tomar nota das entidades que lidam com esta informação bem como os sistemas que a gerem.

Detalhados os primeiros aspetos, passa-se à avaliação dos problemas que deve ser feita ao sistema. O primeiro é a **perda de confidencialidade** visto que os recursos humanos têm que partilhar, por exemplo, os vencimentos dos seus empregados. Como é expectável, essa partilha está sujeita a ataques e a forma como os mesmos são manuseados faz com que este problema tenha uma classificação *MEDIUM*. O segundo é a **perda de integridade e disponibilidade** - pelo facto dos titulares dos dados necessitarem de reenviar/alterar informação ou de haver indisponibilidade na receção do salário, por exemplo. Apesar de tudo, o problema é classificado com o grau de gravidade *LOW*. Para finalizar esta avaliação, importa referir ainda a atenção redobrada que é necessário ter no processamento de determinadas informações relativas à saúde dos empregados: aqui a classificação do problema é cotada com *HIGH* pelo facto de serem informações de caráter extremamente sensível.

Feita uma avaliação dos riscos inerentes a este sistema de dados, mostra-se agora a probabilidade de ocorrência de determinadas ameaças:

- internet e recursos técnicos: a probabilidade de ataques neste campo é baixa visto que não há uma conexão constante à internet;
- procedimentos relacionados com os dados dos empregrados: assumindos que os procedimentos são constantemente revistos e atualizados e que são construidos de acordo com a regulamentação da empresa, a probabilidade de ocorrência de qualquer problema é também baixa;
- pessoal envolvido na gestão dos dados dos empregados: nem sempre é feito um treino apropriado deste pessoal e nem sempre se avaliam as suas características psicológicas, uma vez que essas pessoas podem ter interesses maliciosos; assim a probabilidade de ocorrência de problemas neste caso é média;
- setor comercial e escala de processamento: o tipo de atividade que estamos a falar não reúne as condições que costumam atrair *hackers*, salvo casos com interesses muito específicos, algo que se traduz no reduzido número de ataques que se verificam, estabelecendo a classificação como baixa.

Assim como foram explícitos os problemas, vão-se abordar algumas soluções para os mesmos sendo que algumas passam por:

- **gestão de recursos/bens** - o pessoal que gere as plataformas de dados deve ter um treino específico e as versões do *hardware* e do *software* devem estar devidamente registadas para que se possam associar possíveis vulnerabilidades, por exemplo.
- **política de segurança** - as políticas de segurança devem ser constantemente revistas e atualizadas com uma periodicidade adequeada.
- **segurança do/da *server/database*** - as aplicações e as bases de dados devem funcionar com contas diferentes e os sistemas operativos devem conferir aos seus utilizadores o mínimo de previlégios possíveis.
- **cópias de segurança** - devem ser feitas cópias de segurança regulares dos dados.
- **segurança física** - o perímetro físico relativo às infraestruturas deveria sempre ter um acesso controlado.
- **controlos de acesso e autenticações** - os acessos/alterações/remoções aos dados devem ser registados para que seja possível rastrear as ações que o pessoal executa.