

Mestrado em Engenharia Informática
Universidade do Minho

Engenharia de Segurança

Aula 06 TP - 23/03/2020

João Miranda - PG41845
Sandro Cruz - PG41906

30 de Março de 2020

Conteúdo

1	RGPD (Regulamento Geral de Proteção de Dados)	2
1.1	Experiência 1.1	2
1.2	Pergunta 1.1	3
1.3	Experiência 1.2	5
	1.3.1 Panoptick da Electronic Frontier Foundation (EFF) . .	6
1.4	Experiência 1.3	6
	1.4.1 Exercício 1	6
	1.4.2 Exercício 2	7
	1.4.3 Exercício 3	7
1.5	Experiência 1.4	7
1.6	Pergunta P1.2	7
	1.6.1 Order and delivery of goods	7

Capítulo 1

RGPD (Regulamento Geral de Proteção de Dados)

1.1 Experiência 1.1

RGPD

O RGPD influencia directamente a maneira como é desenvolvido o software, dado que logo na fase de idealização e definição da arquitectura tem que se ter em conta o RGPD.

Com o RGPD o tratamento de dados tem que ser feito de forma mais cuidadosa, todos os dados recolhidos tem que ser explicitamente indicados e só podem ser utilizados para os fins que foram obtidos, e apenas os dados críticos para o software devem ser armazenados durante longos períodos de tempo, a menos que esses dados sejam necessários para fins de arquivo de interesse público, ou para fins de investigação científica ou histórica ou para fins estatísticos. Os dados armazenados deve também ser protegidos contra o tratamento não autorizado ou ilícito e contra a sua perda, destruição ou danificação accidental.

Um responsável deversa definir quais são os dados críticos e os não críticos e deversa definir um "tempo de vida" para este dados, ele deversa também definir quem tem acesso aos dados e quem não tem, sendo que o acesso aos dados só deversa ser dado com intervenção humana e a um conjunto pequeno de indivíduos. A escolha desse responsável deversa ser feita baseando-se nos conhecimentos técnicos na protecção de dados, pelo seu domínio do direito, sendo que a sua função será sempre me agir em conformidade com a lei.

Os dados armazenados deversão também ser pseudonimizados e cifrados, também deversa ser possível assegurar a confidencialidade, integridade, disponibilidade e resiliência permanentes dos sistemas e dos serviços de tratamento, de forma a que esses dados estejam disponíveis caso um ataque físico ao sistema aconteça. O que causa implicações nas bases de dados e como o tratamento e armazenamento dos dados armazenados será feito, dado que será necessária uma boa

definição entre os dados excessivos e os dados críticos, também deveria ser feito um estudo em relação a onde e como serão armazenados os dados, pois será necessário ter sempre disponibilidade dos dados e os dados deverão ser cifrados, protegidos e separados de uma ligação directa a quem é detentor desses mesmos dados.

1.2 Pergunta 1.1

A ENISA produziu o documento *Recommendations on shaping technology according to GDPR provisions - An overview on data pseudonymisation*, sobre o qual se apresenta, posteriormente, um resumo da secção 3: *Pseudonymisation techniques*. A pseudonimização pode contribuir para ocultar a identidade real de um indivíduo, bem como para apoiar a desvinculação em diferentes domínios de processamento de dados. Assim, ao examinar diferentes técnicas de pseudonimização, é importante avaliar se os propósitos acima mencionados podem ser alcançados e em que medida. Uma abordagem de pseudonimização pode gerar ganhos adicionais de proteção de dados em termos de precisão.

De seguida vão ser apresentadas algumas técnicas que podem ser utilizadas para a pseudonimização.

- ***Hashing without key***

Hashing é uma técnica que pode ser usada para suportar a precisão dos dados e derivar pseudónimos. Uma função de *hash* criptográfica h é uma função com propriedades específicas, que transforma qualquer mensagem de entrada, m , de comprimento arbitrário, num *output* de tamanho fixo $h(m)$, designado por valor de *hash* ou *message digest*. Tendo em conta as propriedades de resistência de pré-imagem do *message digest*, observa-se que, para qualquer *hash*, o mesmo *digest* é sempre produzido para o mesmo *input* (bloco de dados).

As funções de *hash* MD5 e SHA-1 devem ser evitadas pois possuem vulnerabilidades conhecidas como a probabilidade de existirem colisões. Portanto, as funções de *hash* criptograficamente resistentes devem ser preferíveis (SHA-2, SHA-3).

As propriedades mencionadas acima permitem que funções de *hash* sejam usadas em várias aplicações, incluindo integridade de dados e autenticação de entidades. No entanto, o *hash* de identificadores de dados para fornecer pseudónimos apresenta enormes desvantagens tais como o facto de qualquer terceiro aplicar a mesma função de *hash* ao mesmo identificador poder obter o mesmo pseudónimo e qualquer terceiro conseguir verificar se um pseudónimo corresponde a um dado identificador.

- ***Hashing with key or salt***

A principal diferença das funções anteriores é que, para o mesmo *input*,

vários pseudónimos diferentes podem ser produzidos, de acordo com a escolha da chave específica. O controlador de dados deve manter a chave secreta armazenada separadamente de outros dados, pois esta fornece os meios para associar os indivíduos aos identificadores originais.

As funções de *hash* com chave podem conduzir ao anónimo de dados, pois o apagamento da chave secreta inibe quaisquer associações entre pseudónimos e os identificadores iniciais. Mais concretamente, tal seria equivalente a gerar um pseudónimo aleatório, sem qualquer conexão aos identificadores iniciais.

Poderiam ser utilizadas funções de *hash* sem chave, mas que recebessem um *salt* que corresponde a um valor aleatório pois, se o este fosse destruído de forma segura pelo controlador, não seria fácil restaurar a associação entre pseudónimos e identificadores.

No entanto, a utilização de *salt* para proteção de *hashes* tem alguns inconvenientes: o *salt* não possui as mesmas propriedades de imprevisibilidade que as chaves secretas (menor tamanho) e as funções de *hash* com chave são consideradas, por norma, criptograficamente mais fortes.

- ***Encryption as a pseudonymisation technique***

A criptografia simétrica dos identificadores dos participantes de dados é também um método eficiente para obter pseudónimos. Nesta, o identificador original de um sujeito de dados pode ser cifrado através de um algoritmo de criptografia simétrica (por exemplo, o AES), fornecendo assim um criptograma que deve ser usado como um pseudónimo, onde a mesma chave secreta é necessária para a decifragem.

A principal diferença desta técnica em relação às funções de *hash* com chave, em termos de pseudonimização, é que o controlador de dados pode sempre obter os identificadores iniciais dos sujeitos de dados, recorrendo à decifragem. No entanto, existem também propriedades idênticas às funções de *hash* com chave tais como:

- A mesma chave secreta deve ser utilizada para fornecer o mesmo pseudónimo para o mesmo identificador.
- Caso a chave seja destruída, não será fácil associar um pseudónimo ao identificador inicial, mesmo que este esteja a ser armazenado pelo controlador de dados.

Deste modo, a criptografia simétrica pode ser aplicada em casos em que um controlador de dados precisa de rastrear os dados, mas também de conhecer os seus identificadores iniciais. Os algoritmos de criptografia de chave pública podem ser usados em casos específicos para fins de pseudonimização.

- ***Other cryptography-based techniques***

A combinação apropriada de vários esquemas criptográficos também pode

fornecer abordagens robustas de pseudonimização, por exemplo, através do uso de técnicas como computação multipartidária segura e criptografia homomórfica. Estas abordagens, embora dispendiosas, parecem ser as melhores opções nos casos em que o princípio da proteção de dados necessita de assegurar que o controlador não deve ter conhecimento prévio da identidade do sujeito de dados, a menos que este opte por provar a sua identidade.

- ***Tokenisation***

Tokenização refere-se ao processo em que os identificadores de indivíduos são substituídos por valores gerados aleatoriamente, conhecidos como *tokens*, sem existir qualquer relação matemática com os identificadores originais. Assim, o conhecimento de um *token* é útil apenas para o controlador ou o processador. No entanto, apesar da eficiência da *tokenização*, a sua implementação pode, dependendo do contexto, ser muito desafiante e complexa. A sincronização de *tokens* em vários sistemas pode ser necessária em várias aplicações.

Concluindo, abordagens que empregam funções de *hash* com chave ou algoritmos de criptografia podem ser preferíveis com relação à redução da complexidade e do armazenamento.

- **Outras abordagens**

Existem muitas outras técnicas conhecidas no contexto de pseudonimização, como *masking*, *scrambling* e *blurring*. *Masking* refere-se ao processo de ocultar parte do identificador de um indivíduo com caracteres aleatórios ou outros dados. *Scrambling* é um processo que pode ser reversível ou não e refere-se a técnicas para misturar ou ofuscar características. *Blurring* é outra técnica, que visa utilizar uma aproximação de valores de dados, de modo a reduzir a precisão dos mesmos, reduzindo a possibilidade de identificação dos indivíduos. Existem também as técnicas associadas aos códigos de barras, códigos QR ou similares, que visam principalmente suportar a precisão dos dados, invés de fornecer uma solução de pseudonimização.

1.3 Experiência 1.2

Neste experiência somente o primeiro e segundo tópicos envolveram uma resposta mais concisa. Os restantes foram verificados com base na visualização e percepção do conteúdo disposto em cada *website*. A resposta do segundo tópico (**PRISM Break**) foi guardada num ficheiro em formato pdf denominado **Windows_PRISMBreak.pdf**.

1.3.1 Panopticlick da Electronic Frontier Foundation (EFF)

Os resultados obtidos da verificação da segurança do *browser* (*Google Chrome*) contra o *tracking* apresentou resultados mistos: existe alguma proteção contra o rastreamento da *web* mas, existem algumas lacunas. É sugerida uma reconfiguração do *software* de proteção ou uma consideração na instalação do *Privacy Badger* da EFF.






Test	Result
Is your browser blocking tracking ads?	 partial protection
Is your browser blocking invisible trackers?	 partial protection
Does your blocker stop trackers that are included in the so-called "acceptable ads" whitelist?	 yes
Does your browser unblock 3rd parties that promise to honor Do Not Track?	 no
Does your browser protect from fingerprinting?	 your browser has a unique fingerprint

Figura 1.1: Resultados da verificação da segurança do *browser* contra o *tracking*.

1.4 Experiência 1.3

1.4.1 Exercício 1

Os nove critérios de avaliação são os seguintes:

- Avaliação ou pontuação
- Tomada de decisão automatizada com efeito legal ou similar significativo
- Monitorização sistemática
- Dados sensíveis ou dados de natureza altamente pessoal
- Dados processados em grande escala
- Correspondência ou combinação de conjuntos de dados
- Dados relativos a sujeitos vulneráveis
- Uso inovador ou aplicação de novas soluções tecnológicas ou organizacionais
- Quando o tratamento em si “impede que os titulares de dados exerçam um direito ou utilizem um serviço ou contrato”

1.4.2 Exercício 2

A título de exemplo consideremos um projeto que consiste numa aplicação de orientação médica para pacientes poderem marcar a sua consulta. Esta aplicação permite o agendamento de consultas, bem como ter acesso ao histórico das consultas pendentes, canceladas ou realizadas, bem como as prescrições médicas de modo a impulsionar uma melhoria no desenvolvimento do processo médico e de vida dos pacientes.

As credenciais de acesso dos pacientes são disponibilizadas pelos seus médicos de família, que aderem à presente aplicação, e as credenciais de acesso dos especialistas são disponibilizadas pelos mesmos.

Cada paciente terá o seu perfil, acompanhado de informações básicas como: nome completo, morada, identificação completa do cartão de cidadão (número, NIF, número de utente de saúde), idade, sexo, doenças, nome de utilizador, endereço de email, foto, password. Cada especialista será caracterizado por: nome completo, nome de utilizador, endereço de email e password.

Após a autenticação dos pacientes, estes serão abordados com algumas questões, tal como se pretende marcar alguma consulta, se pretende ver o histórico e a medicação recomendada.

Todos estes testes serão armazenados de acordo com o nome do paciente, morada, idade, sexo, ano decorrente e especialista responsável pelo seu acompanhamento. Os testes são analisados e acrescentados ao respetivo histórico. Este, por sua vez, apenas está acessível ao paciente e ao seu especialista, tendo este último ainda acesso aos relatórios e medicação.

O paciente pode optar por eliminar a conta solicitando pedido ao especialista, sendo que o histórico anónimo é conservado para fins estatísticos.

1.4.3 Exercício 3

A resolução deste exercício encontra-se num ficheiro pdf em anexo denominado **DPIA_Experiencial1.3.pdf**.

1.5 Experiência 1.4

A resolução deste exercício encontra-se num ficheiro pdf em anexo denominado **DPIA_Experiencial1.4.pdf**.

1.6 Pergunta P1.2

1.6.1 Order and delivery of goods

Pelo que podemos analisar o primeiro passo passara por analisar os diferentes processos que envolvem a encomenda e entrada de produtos, de forma a perceber como o processo poderá ser processado, tratado, entidades envolvidas e

dados que são importantes no trabalho.

Avaliação de Impacto

Nas avaliações de impacto do serviço no caso de perda de confidencialidade e integridade conclui-se que o impacto seria de uma gravidade média, dado que mesmo que haja fuga ou alteração dos dados, mesmos os confidenciais como dados de cartões de crédito, podem só causar uma inconveniência para o detentor dos dados. Apesar que nós achamos que o impacto de uma fuga ou alteração de dados, especialmente de dados confidenciais como números de cartões de crédito, deveriam ter um impacto de gravidade alto, apesar que o documento especifica que num caso geral o impacto poderá variar de gravidade.

Numa avaliação ao impacto causado pela perda da disponibilidade do serviço foi concluído ser baixa, dado que os dados referentes ao serviço poderão ser tratados mais tarde em caso de não estarem disponíveis naquele momento causando impacto apenas na data de entrega do produto.

Então poderá-se concluir que a avaliação de impacto tem uma gravidade média. Apesar que isto apenas se aplica a um caso generalizado de como é normalmente aplicado este tipo de serviço e as gravidades de cada processo poderão variar dependendo do caso.

Probabilidade de Ocorrência de uma Ameaça

- **Recurso Técnicos e Recursos de Rede** - Dado que os dados pessoais presentes na encomenda tem que ser processados pela internet e o sistema que processa a encomenda está ligado a outros sistemas internos ou externos, conclui-se que o risco da ocorrência de uma ameaça é médio. Sendo que se supõe que o acesso não autorizado aos dados é bloqueado pela aplicação das melhores práticas de proteção de dados.
- **Processos Referentes ao Processamento de Dados Pessoais** - A probabilidade de ocorrência é considerada baixa, dado que se supõe que as funções e responsabilidades dos empregados estão bem definidas e vão de acordo com uma política de utilização, os dados apenas são utilizados para os efeitos de processamento da encomenda e existem logs de todas as operações feitas.
- **Pessoa/Entidades Envolvidas no Processamento de Dados** - A probabilidade de ocorrência é considerada média, dado que nem todos os trabalhadores que vão lidar com os dados terão saberão como lidar com os dados.
- **Setor comercial e escala de processamento** - A probabilidade de ocorrência é considerada média, pois uma loja online é propícia a ataques e como as operações relacionadas com uma loja online podem afectar muitos utilizadores a gravidade é considerada média. Também é considerado que

em caso de fuga de dados, esses dados forma tratados de forma a causar o mínimo impacto possível, encriptação e anonimização dos dados.

Avaliação de Risco

Em geral como se pode concluir de risco do impacto de um ataque é médio, como se pode verificar utilizando os dados acima.

Medidas de Mitigação

Como medidas de mitigação de risco sugerimos as seguintes:

- A.3 - Security policy and procedures for the protection of personal data.
- F.4 - Data processors.
- G.3 - Incidents handling / Personal data breaches.
- H.3 - Business continuity.
- I.2 - Confidentiality of personnel.
- J.2 - Training.
- L.5 - Logging and monitoring.
- O.3/O.2 - Network/Communication security.
- M.3 - Server/Database security.