Grupo 9 – Autismo

Bruna Poso - 02.201.007	**Luan Brito - 02.201.030**
Fernando Fialho - 01.201.027	**Luiza Bezerra - 02.201.064**
Jennifer Viana - 02.201.043	**Rodrigo Busto - 02.201.048**

Especificação da implementação de LGPD

Disponibilidades

- Backups seguros de arquivos e dados:

  Para conseguirmos aderir o LGPD de Segurança dos dados pessoais armazenados pensamos em realizar um backup em outra cloud (Azure) do S3 cliente. Para assegurar a confidencialidade dos dados realizaríamos uma criptografia no pyspark da AWS antes de enviar para a Azure, dessa forma os dados estariam seguros e com uma redundância.

  Como funciona: a gente pode utilizar o AzCopy, que vai fazer esse envio de arquivos de uma cloud para outra. Quando inserirmos um arquivo no S3 cliente, automaticamente ele faz uma cópia para a Azure. Porém para evitarmos arquivos maliciosos, dados corrompidos e infecção dos dados temos algumas alternativas de backup:

  - Esses backups seriam D-1 (dia anterior) para evitar que esses dados não confiáveis estejam entrando no banco principal e backup ao mesmo tempo.
  - Os dados finais do cliente terão ao menos 3 cópias seguras, já criptografadas, 2 dessas 3 armazenadas como backup em 2 mídias/servidores diferentes e a outra armazenada em outra cloud (Azure). 1 cópia local adicional também será feita. Semelhante a regra 3-2-1-1-0.

Confidencialidade

- Criptografia dos dados
  - Por padrão o serviço de armazenamento da Azure proporciona criptografia **AES de 256bits**. Também é possível habilitar a dupla criptografia, nos painéis do cloud Azure.

    A proposta de criptografia gira em torno de utilizar o pyspark para criptografar os arquivos gerados que são enviados ao AWS S3, utilizando a biblioteca **cryptography** e usá-la para criar as chaves de encriptação.

    Outra possibilidade é criptografar colunas sensíveis do banco de dados utilizando a funcionalidade de criptografia T-SQL, usando comandos de encriptação.

Grupo 9 – Autismo

Bruna Poso - 02.201.007	**Luan Brito - 02.201.030**
Fernando Fialho - 01.201.027	**Luiza Bezerra - 02.201.064**
Jennifer Viana - 02.201.043	**Rodrigo Busto - 02.201.048**

- Anonimizar dados sensíveis

  O mascaramento de dados (termo que engloba anonimização, pseudo-anonimização, redefinição, limpeza ou desidentificação de dados) é um método de proteção de dados sensíveis que substitui o valor original por um valor equivalente fictício, mas realista. Esse método também pode ser chamado de camuflagem de dados.

  Além disso, dentro desse conceito nós temos alguns métodos e técnicas para aplicá-lo, dentre os métodos:

  Dentre as técnicas:

  - Criptografia: codifica os dados a partir de cálculos matemáticos e algoritmos. Usado quando é necessário voltar os dados para o seu estado original
  - Substituição: os dados são substituídos por outro valor, podendo variar bastante a forma como pode ser feita essa substituição.
  - Descarte: descartar/excluir dados pessoais que não vamos utilizar, dessa forma garantindo a segurança dos dados e a confiabilidade.

Integridade

- Controle de acesso e envio de informação

  Para ter mais precisão nos dados existiria um código para controle de acesso do smartwatch.

  Como funciona: quando o usuário for utilizar o seu smartwatch precisará colocar um código de confirmação no seu aparelho físico ou uma senha numérica, afirmando que é ele mesmo que está portando o aparelho, assim começando o envio de dados para a nuvem. Em caso de erro ou valores aparecerem como "nulos" (indicando que ele não está mais utilizando o smartwatch) também será solicitado que o usuário insira seu código novamente, para que o envio de dados se inicie.

  Motivo: implantação para evitar outra pessoa utilizar o smartwatch e os dados não serem contabilizados, assim tendo mais precisão na análise e envio dos dados do paciente.