GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29.º

1021/00/PT
WP207

Parecer 06/2013 sobre os dados abertos e a reutilização de informações do setor público («ISP»)

Adotado em 5 de junho de 2013

O GRUPO DE PROTEÇÃO DAS PESSOAS SINGULARES NO QUE DIZ RESPEITO AO TRATAMENTO DE DADOS PESSOAIS

instituído pela Diretiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995,

tendo em conta os artigos 29.º e 30.º, n.º 1, alínea a*)*, e n.º 3 da referida Diretiva,

tendo em conta o seu regulamento interno,

ADOTOU O PRESENTE PARECER:

I. Introdução

1.1. Revisão da Diretiva ISP

Em 26 de junho de 2013, a União Europeia adotou a Diretiva 2013/37/UE do Parlamento Europeu e do Conselho (a «Alteração ISP»), que altera a Diretiva 2003/98/CE relativa à reutilização de informações do setor público (a «Diretiva ISP»)[1].
A Diretiva ISP visa facilitar a reutilização de informações do setor público através da harmonização das condições de reutilização em toda a União Europeia e da eliminação de obstáculos desnecessários à reutilização no mercado interno.

O texto inicial da Diretiva ISP, adotado em 2003, harmonizava as condições de reutilização, mas não impunha aos organismos do setor público a obrigação de disponibilizar dados para reutilização. A disponibilização de dados para reutilização era essencialmente facultativa: esta decisão era deixada ao critério dos Estados-Membros e dos organismos do setor público em causa. Consequentemente, muitos organismos do setor público europeus optavam simplesmente por não autorizar a reutilização das suas informações.

Em face deste cenário, um dos principais objetivos da Alteração ISP consiste em estabelecer o princípio de que todas as informações públicas (ou seja, todas as informações na posse do setor público que estejam disponíveis ao público nos termos da legislação nacional) são reutilizáveis para fins comerciais e não comerciais. Em certos casos, existem exceções ao âmbito de aplicação da Diretiva ISP alterada, nomeadamente por motivos de proteção de dados[2].

Assim, a Diretiva ISP alterada impõe agora sobre os organismos do setor público a obrigação de autorizar a reutilização de todas as informações públicas na sua posse. No entanto, tal como será demonstrado mais adiante, não obriga os organismos do setor público a divulgarem publicamente informações pessoais. Só estabelece a obrigação de tornar reutilizáveis as informações que estejam já disponíveis ao público nos termos da legislação nacional e, mesmo nesse caso, apenas se a reutilização não contrariar as disposições da legislação sobre proteção de dados aplicável.

Outras disposições relevantes da Alteração ISP alargam o âmbito de aplicação da Diretiva ISP às bibliotecas (incluindo as bibliotecas universitárias), aos arquivos e aos museus.

À luz do exposto, a Diretiva ISP alterada tem o potencial para aumentar consideravelmente a acessibilidade das informações na posse dos organismos públicos.

1 JO L 175 de 27.6.2013, p. 1.

2 A questão do âmbito de aplicação da Diretiva ISP alterada e das disposições relacionadas com a proteção de dados é desenvolvida na secção V.

1.2. Reutilização de ISP e dados pessoais

As iniciativas de reutilização de ISP envolvem geralmente (i) a disponibilização de bases de dados completas (ii) em formato eletrónico normalizado, (iii) a qualquer interessado, sem processo de controlo prévio, (iv) gratuitamente (ou mediante o pagamento de taxas de valor reduzido), e (v) para quaisquer fins comerciais ou não comerciais, sem condições (ou, em certos casos, sob condições não restritivas, através de uma licença)[3].

Embora estas iniciativas possam trazer benefícios que conduzam a maior transparência e à reutilização inovadora das informações do setor público, a maior acessibilidade das informações daí resultante comporta riscos.

A fim de minimizar estes riscos, sempre que estejam envolvidos dados pessoais, a legislação sobre proteção de dados deve ajudar a orientar o processo de seleção dos dados pessoais que podem ser disponibilizados para reutilização e das medidas a adotar para salvaguardar estes dados. Sempre que esteja em causa a proteção da privacidade e dos dados pessoais, é necessário seguir uma abordagem equilibrada. Por um lado, as regras que visam a proteção dos dados pessoais não devem constituir um obstáculo injustificado ao desenvolvimento do mercado da reutilização. Por outro lado, o direito à proteção dos dados pessoais e o direito à privacidade têm de ser respeitados. Importa salientar que o conceito de «dados abertos» visa, antes de mais, promover a transparência e a responsabilização dos organismos do setor público, bem como o crescimento económico, e não a transparência dos cidadãos.

Ao aplicarem a Diretiva ISP e a legislação sobre proteção de dados à reutilização de dados pessoais, os organismos do setor público tomarão provavelmente uma das três seguintes decisões:

1. Decisão de não disponibilizar informações pessoais para reutilização ao abrigo da Diretiva ISP;
2. Decisão de anonimizar as informações pessoais (convertendo-as normalmente em dados estatísticos agregados)[4] e de disponibilizar para reutilização apenas esses dados anonimizados;
3. Decisão de disponibilizar informações pessoais para reutilização (quando necessário, estabelecendo condições específicas e garantias adequadas).

II. Objetivo do parecer

2.1. Orientações consistentes e melhores práticas

O presente parecer tem por objetivo ajudar a assegurar um entendimento comum sobre o quadro jurídico aplicável e proporcionar orientações consistentes e exemplos de melhores práticas sobre a aplicação da Diretiva ISP (na redação atualmente em vigor) no que respeita ao tratamento de dados pessoais.

Por conseguinte, não pretende harmonizar as abordagens nacionais quanto ao nível de transparência, nem a legislação nacional sobre o acesso aos documentos ou a disponibilidade das informações ao abrigo dessa legislação. Contudo, as divergências na legislação nacional que transpõe a Diretiva ISP

[3] Note-se que, de acordo com o artigo 8.º, n.º 1, da Diretiva ISP, na redação atualmente em vigor, as condições da licença «não devem restringir desnecessariamente as possibilidades de reutilização e não devem ser utilizadas para limitar a concorrência».

[4] A questão da reutilização de conjuntos de dados agregados e anonimizados obtidos a partir de dados pessoais é desenvolvida na secção V.

e na interpretação nacional da Diretiva 95/46/CE[5] no que respeita à reutilização de ISP nem sempre podem ser explicadas pela necessidade de ter em conta a diversidade dos regimes nacionais de acesso e os diferentes níveis de transparência.

Nesta matéria, as Recomendações Estratégicas sobre Privacidade, emitidas pela rede temática LAPSI em setembro de 2012, ilustram claramente as disparidades desnecessárias na forma como a Diretiva ISP foi transposta nos Estados-Membros no que respeita à proteção dos dados pessoais[6]. A própria Diretiva ISP alerta para a possibilidade de essas diferenças e incertezas no plano legislativo se tornarem mais significativas com o desenvolvimento da sociedade da informação, que conduziu já a um grande aumento da exploração transfronteiriça da informação[7].

A ausência de uma abordagem consistente poderá fragilizar a posição das pessoas em causa. Poderá igualmente impor encargos regulamentares desnecessários sobre as empresas e outras organizações com atividade transfronteiriça e, como tal, representar um obstáculo ao desenvolvimento de um mercado europeu comum para a reutilização. Por um lado, é necessário fornecer às pessoas em causa garantias de que os seus dados serão consistentemente protegidos, independentemente da sua transferência para outro Estado-Membro para fins de reutilização. Por outro lado, importa também evitar uma complexidade e fragmentação injustificadas para permitir a livre circulação de dados pessoais na Europa, que constitui outro dos principais objetivos da Diretiva 95/46/CE.

2.2. Necessidade de atualizar o Parecer 7/2003

A Alteração ISP tem lugar uma década após a adoção da Diretiva ISP em 2003. Naquela altura, o GT29 adotou um parecer sobre questões suscitadas pelas ISP em matéria de proteção de dados («Parecer 7/2003»)[8]. Embora os princípios fundamentais enunciados no Parecer 7/2003 permaneçam válidos, a evolução tecnológica e outros desenvolvimentos no campo das ISP e da proteção de dados, incluindo as alterações legislativas propostas em ambos os campos, justificam os atuais esforços para atualizar e complementar o parecer de 2003.

Além disso, o parecer pode agora tomar também em consideração outros esforços recentes e ainda em curso que visam proporcionar orientações adicionais, em especial:

- o parecer da Autoridade Europeia para a Proteção de Dados («AEPD»), de 18 de abril de 2012, sobre o «Pacote dados abertos» da Comissão[9];
- o Parecer 3/2013 do GT29 sobre a limitação da finalidade[10];

---

[5] Diretiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995, relativa à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados (JO L 281 de 23.11.1995, p. 31).

[6] Fundada pela Comissão Europeia, a LAPSI é uma rede temática europeia dedicada ao tema dos Aspetos Jurídicos das Informações do Setor Público; ver http://www.lapsi-project.eu/. As recomendações estão disponíveis em http://www.lapsi-project.eu/lapsifiles/lapsi_privacy_policy.pdf.

[7] Ver o considerando 7.

[8] Ver o Parecer 7/2003 do Grupo de Trabalho de Proteção de Dados do artigo 29.º sobre a reutilização da informação do setor público e proteção de dados pessoais – Estabelecer um equilíbrio – adotado em 12 de dezembro de 2003 (WP 83). Ver também dois pareceres anteriores do GT29 relacionados com esta matéria: Parecer 3/1999 relativo a informação do setor público e proteção de dados pessoais, adotado em 3 de maio de 1999 (WP20), bem como o Parecer 5/2001 sobre um relatório especial do Provedor de Justiça Europeu, adotado em 17 de maio de 2001.

[9] Parecer da AEPD sobre o «Pacote dados abertos» da Comissão Europeia, que inclui uma Proposta de Diretiva do Parlamento Europeu e do Conselho que altera a Diretiva 2003/98/CE relativa à reutilização de informações do setor público (ISP), uma Comunicação da Comissão sobre dados abertos e a Decisão da Comissão 2011/833/UE relativa à reutilização de documentos da Comissão, adotado em 18 de abril de 2012, disponível em: http://www.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-04-18_Open_data_EN.pdf.

- o trabalho atualmente em curso do subgrupo «Tecnologia» do GT29 sobre técnicas de anonimização[11];
- o trabalho desenvolvido em alguns Estados-Membros sobre anonimização e avaliação dos riscos[12]; e
- a atual jurisprudência e prática sobre o equilíbrio entre a reutilização e a proteção de dados pessoais em alguns Estados-Membros[13].

III. Objeto e estrutura do parecer

O Parecer 7/2003 centrava-se no princípio da limitação da finalidade[14], mas também abordava outras questões, tais como a legitimidade da comunicação pública e da reutilização das ISP, a proteção especial conferida aos dados sensíveis, as transferências para países terceiros, a qualidade dos dados e os direitos das pessoas em causa. Estas observações ainda são válidas. Tendo em conta o trabalho já realizado anteriormente, o presente parecer limita-se a atualizar e a complementar as conclusões do Parecer 7/2003 sempre que tal se mostre necessário à luz dos desenvolvimentos legislativos e tecnológicos.

A secção IV esclarece que a obrigação de reutilização prevista na Diretiva ISP alterada não prejudica os requisitos em matéria de proteção de dados e salienta a importância da «proteção de dados desde a conceção» e da «proteção de dados por defeito», bem como das «avaliações do impacto na proteção de dados», para assegurar que é dada resposta às questões suscitadas em matéria de proteção de dados antes de os dados pessoais serem disponibilizados para reutilização.

A secção V fornece orientações através de exemplos que ilustram o tipo de dados pessoais suscetíveis de serem abrangidos pelo âmbito de aplicação da Diretiva ISP.

A secção VI é dedicada às situações atualmente mais comuns em iniciativas de reutilização de ISP: a disponibilização de dados estatísticos agregados e anonimizados, obtidos a partir de dados pessoais. É o caso, por exemplo, dos dados estatísticos agregados sobre as taxas de criminalidade, a despesa pública ou o desempenho escolar das crianças em diferentes zonas geográficas ou em diferentes estabelecimentos de ensino. Uma vez que este é o cenário mais comum de reutilização de informações do setor público que contêm dados pessoais, uma parte significativa do presente parecer ser-lhe-á dedicada. Aqui, a principal preocupação em termos de proteção de dados consiste em assegurar uma agregação e anonimização eficaz dos dados e em minimizar o risco de reidentificação de dados pessoais a partir dos conjuntos de dados agregados.

A secção VII analisa, de forma menos pormenorizada, situações em que dados pessoais são disponibilizados ao público e, como tal, poderão estar disponíveis para reutilização. Embora este não seja atualmente o cenário típico das iniciativas de reutilização de ISP, é importante ter em conta que os organismos do setor público disponibilizam, cada vez mais, dados pessoais ao público, muitas vezes na Internet. Frequentemente, estão em causa dados pessoais diretamente identificáveis como, por exemplo, as informações constantes do registo predial sobre o proprietário de um determinado

---

[10] Parecer 3/2013 do Grupo de Trabalho de Proteção de Dados do artigo 29.º sobre a limitação da finalidade, adotado em 2 de abril de 2013 (WP 203).

[11] Está prevista a adoção de um parecer sobre esta matéria no segundo semestre de 2013.

[12] Ver, por exemplo, o código de prática sobre anonimização intitulado «Anonymisation: Managing data protection risk code of practice» (Anonimização: código de prática de gestão dos riscos em matéria de proteção de dados), publicado pelo Gabinete do Comissário para a Informação do Reino Unido em novembro de 2012 e as orientações sobre análise dos riscos publicadas pela autoridade francesa para a proteção de dados em junho de 2012.

[13] Ver, por exemplo, a Recomendação Estratégica da LAPSI de setembro de 2012 (p. 4-14).

[14] Ver o artigo 6.º, n.º 1, alínea b), da Diretiva 95/46/CE.

imóvel, as declarações de interesses ou os salários de certos funcionários públicos ou as despesas dos deputados. Coloca-se aqui a questão de saber em que medida, para que finalidades, em que condições e com que garantias estes dados podem ser disponibilizados para reutilização. Importa igualmente esclarecer se estes dados são abrangidos pelas disposições da Diretiva ISP.

Neste contexto, é importante salientar que quaisquer informações relativas a uma pessoa singular identificada ou identificável, estejam ou não publicamente disponíveis, constituem dados pessoais. Por conseguinte, o acesso a dados pessoais que tenham sido disponibilizados ao público e a sua reutilização (por exemplo, publicando os dados na Internet) continuam sujeitos à legislação sobre proteção de dados aplicável.

Alguns cenários específicos, como o caso dos dados da investigação e a situação dos arquivos históricos – que estão agora abrangidos pelo âmbito de aplicação da Diretiva ISP – serão sucintamente abordados nas secções VIII e IX.

A secção IX analisa a questão do licenciamento de ISP e a necessidade de incluir uma cláusula sobre proteção de dados nas licenças, sempre que tal for relevante.

Por último, a secção XI apresenta um conjunto de conclusões e recomendações.

IV. Nem todos os dados pessoais «publicamente disponíveis» devem ser disponibilizados para reutilização

4.1. A obrigação de reutilização prevista na Diretiva ISP não prejudica os requisitos em matéria de proteção de dados

Quando foi adotada em 2003, a Diretiva ISP não impunha aos organismos do setor público a obrigação de autorizarem a reutilização de ISP. A decisão de autorizar ou não essa reutilização continuava a caber aos Estados-Membros ou ao organismo do setor público em causa (de acordo com o quadro regulamentar nacional sobre transparência e acesso). O Parecer 7/2003 foi adotado à luz desta «não-obrigação». A secção 2(cc) do Parecer 7/2003 afirma que «É importante sublinhar que a diretiva relativa à reutilização não pode ser invocada como obrigação legal a cumprir, uma vez que não cria a obrigatoriedade de comunicar dados pessoais».

Com a Alteração ISP, a análise torna-se mais complexa, mas a conclusão final é a mesma.

O artigo 3.º, n.º 1, da Diretiva ISP alterada estabelece que «sem prejuízo do disposto no n.º 2, os Estados-Membros devem assegurar que os documentos aos quais a presente diretiva é aplicável nos termos do n.º 1 sejam reutilizáveis para fins comerciais ou não comerciais, de acordo com as condições previstas nos capítulos III e IV.» A menos que seja possível recusar a reutilização pelos motivos previstos no artigo 1.º (motivos resultantes dos regimes nacionais de acesso e também especificamente por razões de proteção dos dados pessoais), a reutilização tem de ser autorizada.

Simultaneamente, o considerando 21 da Diretiva ISP refere que esta «deve ser aplicada e executada no pleno cumprimento dos princípios relativos à proteção de dados pessoais». Além disso, o artigo 1.º, n.º 4, estabelece que a Diretiva ISP «não modifica, nem de modo algum afeta o nível de proteção dos indivíduos relativamente ao processamento de dados pessoais».

Estas disposições, lidas em conjunto, significam que o «princípio da reutilização» não é de aplicação automática quando esteja em causa o direito à proteção dos dados pessoais e não prevalece sobre as disposições aplicáveis da legislação sobre proteção de dados. Quando documentos existentes na posse de organismos do setor público contenham dados pessoais, a sua reutilização cai no âmbito de

aplicação da Diretiva 95/46/CE e, como tal, continua a estar sujeita à legislação sobre proteção de dados aplicável.

Consequentemente, nos casos em que a reutilização abranja dados pessoais, o organismo do setor público não pode invocar sistematicamente a necessidade de cumprir a Diretiva ISP como fundamento legítimo para disponibilizar os dados para reutilização.[15]

4.2. Importância da avaliação do impacto na proteção de dados antes de disponibilizar os dados para reutilização

Tendo em conta os potenciais riscos da reutilização de ISP e, em especial, o facto de ser muito difícil controlar eficazmente a utilização dos dados pessoais que tenham sido disponibilizados ao público para reutilização, o GT29 reitera a necessidade de respeitar os princípios da «proteção de dados desde a conceção» e da «proteção de dados por defeito» e de dar resposta às questões suscitadas em matéria de proteção de dados numa fase precoce. Em especial, o GT29 recomenda fortemente a realização de uma avaliação minuciosa do impacto na proteção de dados pelo organismo do setor público antes de disponibilizar publicamente dados pessoais para reutilização. Os Estados-Membros devem igualmente estudar a possibilidade de estabelecer a obrigatoriedade legal desta avaliação de impacto ou de a promover como boa prática. Em qualquer caso, mesmo que tal não esteja expressamente previsto na legislação nacional, antes da divulgação das informações e da decisão de as disponibilizar para reutilização, os organismos do setor público devem realizar uma avaliação minuciosa para determinar se podem ser disponibilizados dados pessoais para reutilização e, em caso afirmativo, em que condições tal reutilização é admissível e que garantias específicas de proteção de dados devem ser impostas.

Esta avaliação deve, por exemplo, determinar a base jurídica para a divulgação (e a potencial base jurídica para a reutilização), analisar os princípios da limitação da finalidade, da proporcionalidade e da minimização dos dados e ter em consideração a proteção especial a conferir aos dados sensíveis. Durante esta avaliação, importa analisar cuidadosamente o impacto potencial sobre as pessoas em causa.

Esta avaliação deve ajudar a decidir quais os dados pessoais que poderão eventualmente ser disponibilizados para reutilização e quais as garantias a aplicar[16]. Importa salientar que o Regulamento «Proteção de Dados» proposto[17] aconselha e, em alguns casos, exige a realização de avaliações do impacto na proteção de dados, considerando que constituem uma ferramenta fundamental para ajudar a assegurar a responsabilização dos responsáveis pelo tratamento de dados[18].

---

[15] O GT29 deseja também deixar bem claro que, da perspetiva do reutilizador, a Diretiva ISP, por si só, também não cria um fundamento legítimo para o tratamento dos dados. (Relativamente à questão dos fundamentos legítimos, ver o Parecer 7/2003 e a secção 7.5 *infra*.)

[16] Se a avaliação conduzir à decisão de não disponibilizar para reutilização dados pessoais enquanto tal, mas de disponibilizar antes conjuntos de dados anonimizados obtidos a partir de dados pessoais, deve ser realizada uma avaliação do risco de reidentificação. Ver secção VI sobre anonimização e avaliação do risco de reidentificação.

[17] Em 25 de janeiro de 2012, a Comissão adotou um pacote para reformar o quadro europeu da proteção de dados. Este pacote inclui (i) uma «Comunicação» (COM(2012) 9 final), (ii) uma «Proposta de Regulamento Proteção de Dados» (COM(2012) 11 final), e (iii) uma «Proposta de Diretiva Proteção de Dados» (COM(2012) 10 final).

[18] Para mais orientações sobre a realização de uma avaliação do impacto na proteção de dados, ver, por exemplo, o sítio *Web* do projeto PIAF (Um Quadro de Avaliação do Impacto na Privacidade para promover os direitos à proteção de dados e à privacidade) em http://www.piafproject.eu/Index.html. O PIAF é um projeto cofinanciado pela Comissão Europeia que visa encorajar a UE e os seus Estados-Membros a adotar uma política progressiva de avaliação do impacto na privacidade como forma de satisfazer as necessidades e responder aos desafios relacionados com a privacidade e o tratamento de dados pessoais. Também estão disponíveis orientações em alguns Estados-Membros.

Sempre que possível, a análise que precede a decisão de reutilização deve ter por base um debate informado com a participação de várias partes interessadas, incluindo não só o responsável pelo tratamento de dados que pretende publicar os dados como também aqueles que exigem acesso aos dados e que, como tal, podem fornecer o contexto para o debate, bem como os representantes das pessoas cujos dados pessoais estão em causa (por exemplo, organizações de defesa dos consumidores, organizações de defesa dos direitos dos pacientes, sindicatos de professores, etc.). Quando este debate não for totalmente esclarecedor, a autoridade para a proteção de dados competente e as autoridades nacionais responsáveis pela liberdade de informação poderão fornecer orientações.

Os Estados-Membros devem ainda ponderar a criação e o apoio de redes de conhecimento/centros de excelência e viabilizar, deste modo, a partilha de boas práticas relacionadas com a anonimização e os dados abertos. Estas estruturas poderão ser particularmente importantes para os organismos do setor público mais pequenos, que poderão não possuir os conhecimentos necessários para realizar a anonimização e a avaliação do impacto na proteção de dados, nem para avaliar e testar os riscos de reidentificação[19].

Por último, também é fortemente aconselhável realizar uma avaliação de impacto antes da aprovação de nova legislação que exija a divulgação pública de dados pessoais.

V. Âmbito de aplicação da Diretiva ISP: exceções por motivos de proteção de dados pessoais

A presente secção contém orientações sobre o âmbito de aplicação da Diretiva ISP e, em especial, sobre as exceções estabelecidas por motivos de proteção de dados.

5.1. Aplicabilidade do quadro geral da proteção de dados à reutilização de ISP

O considerando 21 da Diretiva ISP refere que a mesma «deve ser aplicada e executada no pleno cumprimento dos princípios relativos à proteção de dados pessoais». Além disso, o artigo 1.º, n.º 4, estabelece que a Diretiva ISP «não modifica, nem de modo algum afeta o nível de proteção dos indivíduos relativamente ao processamento de dados pessoais».

5.2. Exceções por motivos de proteção de dados pessoais

A Diretiva ISP estabelece que «a presente diretiva não é aplicável a: … documentos não acessíveis por força dos regimes de acesso dos Estados-Membros …»[20].

Ver, por exemplo, o manual da avaliação do impacto na privacidade (AIP) publicado pelo Comissário para a Informação do Reino Unido, disponível em: http://ico.org.uk/for_organisations/data_protection/topic_guides/privacy_impact_assessment; as orientações sobre a análise dos riscos publicadas pela autoridade francesa para a proteção de dados, já referidas na nota 12 *supra*; e as orientações fornecidas pelo Comissário para a Informação esloveno, especificamente sobre as avaliações do impacto na privacidade nos projetos de administração em linha, disponíveis em: https://webmail.europarl.europa.eu/exchweb/bin/redir.asp?URL=https://www.ip-rs.si/fileadmin/user_upload/Pdf/smernice/PIASmernice__ENG_Lektorirano_10._6._2011.pdf

[19] Por exemplo, no Reino Unido, um consórcio liderado pela Universidade de Manchester, do qual fazem parte a Universidade de Southampton, o Serviço de Estatísticas Nacionais e o Instituto dos Dados Abertos (ODI) recentemente criado pelo Governo, gere a Rede de Anonimização do Reino Unido (UKAN), que visa facilitar a partilha de boas práticas relacionadas com a anonimização nos setores público e privado. A rede possui um sítio *Web* em https://webmail.europarl.europa.eu/exchweb/bin/redir.asp?URL=http://www.ukanon.net, e inclui estudos de casos, oficinas e seminários.

[20] Ver artigo 1.º, n.º 2, alínea c), da Diretiva ISP.

Além disso, a Diretiva ISP, na redação atualmente em vigor, também prevê exceções por motivos de proteção de dados. O artigo 1.º, n.º 2, alínea c-C*)*, contempla três situações que estão excluídas do âmbito de aplicação da Diretiva ISP:

- documentos não acessíveis por força dos regimes de acesso por motivos de proteção de dados pessoais;
- documentos de acesso restrito por força dos regimes de acesso por motivos de proteção de dados pessoais; e
- «partes de documentos acessíveis por força desses regimes que contêm dados pessoais cuja reutilização foi definida por lei como incompatível com a legislação relativa à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais».

5.3. Observações gerais

O GT29 sublinha que, independentemente do «princípio da reutilização» formulado na Alteração ISP, a reutilização para fins comerciais ou não comerciais ao abrigo da Diretiva ISP nem sempre será adequada nos casos em que as ISP a reutilizar contenham dados pessoais. As decisões sobre a reutilização de dados pessoais ao abrigo da Diretiva ISP terão de ser tomadas caso a caso, sendo ainda necessário adotar medidas jurídicas, técnicas ou organizativas adicionais para proteger as pessoas em causa.

A reutilização de dados pessoais publicamente disponíveis é e deve ser condicionada:

- pelas disposições gerais da legislação sobre proteção de dados aplicável;
- (se for o caso) por restrições jurídicas adicionais específicas; e
- pelas garantias técnicas e organizativas que terão de ser aplicadas para proteger os dados pessoais.

5.4. Documentos não acessíveis

Esta disposição exclui do âmbito de aplicação da Diretiva ISP todos os documentos não acessíveis por força dos regimes de acesso do Estado-Membro em causa por motivos de proteção de dados pessoais.

Ao contrário do que acontece com a legislação sobre proteção de dados, que se encontra harmonizada, em grande parte, com base na Diretiva 95/46/CE, a legislação sobre o acesso à informação varia significativamente entre os Estados-Membros da UE. Por norma, os regimes de acesso aplicam um critério de equilíbrio, comparando os interesses protegidos pelas regras relativas à privacidade e à proteção de dados com os benefícios do livre acesso e da transparência. Tendo em conta as divergências existentes, o resultado desta comparação poderá ser diferente em diferentes Estados-Membros da UE. Por exemplo, as autoridades fiscais de alguns Estados-Membros podem publicar certas partes das declarações do imposto sobre os rendimentos dos contribuintes (sendo aplicáveis medidas jurídicas, técnicas e organizativas para minimizar os riscos de utilização abusiva), enquanto outros Estados-Membros consideram que estas informações são abrangidas pela exceção e, em princípio, não devem ser divulgadas ao público.

Seja como for, a legislação nacional tem de cumprir o artigo 8.º da Convenção Europeia dos Direitos do Homem («CEDH») e os artigos 7.º e 8.º da Carta dos Direitos Fundamentais da União Europeia («Carta da UE»). Por conseguinte, tal como sustentou o Tribunal de Justiça da União Europeia nos

acórdãos *Österreichischer Rundfunk* e *Schecke*[21], é necessário determinar se a divulgação é necessária e proporcional ao objetivo legítimo prosseguido pela lei.

Em qualquer caso, se a legislação do Estado-Membro em causa estabelecer que os dados pessoais contidos num documento não são acessíveis (incluindo as situações em que a legislação nacional sobre transparência e abertura não prevê a acessibilidade geral dos dados pessoais em questão), esses dados estarão igualmente excluídos do âmbito de aplicação da Diretiva ISP.

A fim de garantir a segurança jurídica e a transparência perante as pessoas em causa, é boa prática, sempre que possível, adotar uma abordagem proativa e definir previamente os dados pessoais que poderão ser disponibilizados ao público. No momento da recolha dos dados, as pessoas em causa poderão então ser informadas se alguma parte dos dados pessoais que fornecem, ou que serão objeto de um tratamento posterior durante o processo administrativo, serão disponibilizados ao público por força da legislação sobre liberdade de informação.

5.5. Documentos de acesso restrito

Esta disposição exclui do âmbito de aplicação da Diretiva ISP todos os documentos de acesso restrito por força dos regimes de acesso do Estado-Membro em causa por motivos de proteção de dados pessoais. Mais uma vez, as disposições dos regimes de acesso sobre os dados que poderão ser sujeitos a um acesso restrito e os tipos de restrições aplicáveis poderão variar de Estado-Membro para Estado-Membro. Eis alguns exemplos:

- coleções de arquivos nacionais contendo dados pessoais que estão sujeitos a condições de acesso específicas e a garantias adicionais (ver secção IX *infra*);
- coleções de dados da investigação contendo dados pessoais que estão sujeitos a condições de acesso específicas e a garantias adicionais (ver secção VIII *infra*);
- certas informações constantes de registos públicos, processos judiciais ou outros documentos administrativos contendo dados pessoais aos quais apenas têm acesso as pessoas singulares ou coletivas que demonstrem um interesse legítimo, ou que estão sujeitos a outras condições de acesso específicas e a garantias adicionais.

5.6. Partes de documentos acessíveis mas cuja reutilização é incompatível

Esta disposição exclui do âmbito de aplicação da Diretiva ISP:

- partes de documentos
- acessíveis por força dos regimes de acesso nacionais
- que contêm dados pessoais «cuja reutilização foi definida por lei como incompatível com a legislação relativa à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais».

Esta disposição confirma que, mesmo nos casos em que certos documentos que contêm dados pessoais são plenamente acessíveis, a sua reutilização pode, ainda assim, ser restringida por motivos de proteção de dados.

[21] Ver acórdão do TJUE de 20 de maio de 2003, *Rundfunk*, nos processos apensos C-465/00, C-138/01 e C-139/01 e acórdão do TJUE de 9 de novembro de 2010, *Volker und Markus Schecke*, nos processos apensos C-92/09 e C-93/09.

O GT29 salienta que esta disposição deve ser interpretada à luz do artigo 1.º, n.º 4, da Diretiva ISP, nos termos do qual esta «não modifica, nem de modo algum afeta o nível de proteção dos indivíduos relativamente ao processamento de dados pessoais».

O GT29 considera que seria uma boa prática adotar, a nível nacional, disposições legais específicas que descrevessem claramente (i) os dados que são disponibilizados ao público, (ii) para que fins, e (iii) sendo o caso, em que medida e em que condições é autorizada a reutilização. No entanto, a ausência deste tipo de disposições não significa que os dados pessoais disponibilizados ao público possam ser sempre reutilizados ao abrigo da Diretiva ISP.

Nestes casos, será a legislação sobre proteção de dados (aplicada em conjunto com outra legislação relevante, como a legislação sobre o acesso a documentos) a determinar se, naquela situação concreta, podem ser disponibilizados dados pessoais para reutilização e, em caso afirmativo, quais as garantias adicionais a implementar. Se o resultado desta análise for positivo, a reutilização será autorizada, sujeita a garantias de proteção de dados específicas e a todas as outras condições estabelecidas na Diretiva ISP (desde que não contrariem a legislação sobre proteção de dados). Se o resultado da análise for negativo, a reutilização estará fora do âmbito de aplicação da Diretiva ISP:

Os exemplos que se seguem poderão ajudar a ilustrar os casos em que esta exclusão do âmbito de aplicação da Diretiva ISP é aplicável. No primeiro exemplo, as restrições à reutilização estão claramente estabelecidas na lei.

- a legislação fiscal de um Estado-Membro estabelece que as declarações do imposto sobre o rendimento de todos os residentes daquele país estarão publicamente disponíveis, podendo ser consultadas por qualquer outro residente, mediante requerimento, nas instalações das autoridades fiscais, sem necessidade de demonstrar um interesse legítimo. Essa legislação estabelece também claramente que os dados não podem ser objeto de um tratamento posterior, por exemplo, publicados na Internet, combinados com outros dados ou editados. Uma ONG solicita acesso à base de dados sobre declarações fiscais e autorização para reutilizar esses dados publicando-os no seu sítio *Web*. Neste caso, os dados fiscais estão fora do âmbito de aplicação da Diretiva ISP e o organismo do setor público não está obrigado a disponibilizar o conjunto de dados para reutilização ao abrigo dessa Diretiva.

Em muitos outros casos, porém, é pouco provável que as restrições jurídicas sejam expressas de forma tão clara e categórica em termos de reutilização. Vários registos civis, comerciais e demográficos, bem como outras bases de dados, permitem a consulta de dados pessoais pelo público em geral, cada vez mais em formato digital através da Internet. A acessibilidade está muitas vezes sujeita a garantias específicas, incluindo restrições técnicas às capacidades de pesquisa e ao descarregamento de grandes quantidades de dados. Poderá ser também solicitado aos utilizadores que aceitem termos e condições de acesso.

- a legislação fiscal de um Estado-Membro estabelece que os nomes dos residentes com impostos em atraso acima de um determinado montante há um largo período de tempo serão publicados num sítio *Web* específico, durante um prazo limitado, sendo implementadas garantias técnicas adicionais, nomeadamente limites ao descarregamento de grandes quantidades de dados e às capacidades de pesquisa. Esta publicação visa incentivar o pagamento pontual do imposto sobre o rendimento e aplicar aos infratores uma sanção adicional que afete a sua reputação. Um consórcio de bancos solicita acesso para reutilização, com vista a introduzir os dados no seu sistema de informações sobre crédito.
- a legislação específica aplicável no setor da saúde de um Estado-Membro permite, mediante certas garantias, que os pacientes verifiquem, num sítio *Web* especial, se um determinado

médico ou outro profissional foi proibido de exercer a sua atividade. São aplicáveis garantias técnicas, tais como limites ao descarregamento de grandes quantidade de dados e às capacidades de pesquisa. Uma organização de defesa dos direitos dos pacientes solicita acesso para reutilização, tendo em vista a criação de um sítio *Web* multilingue e de utilização mais fácil para aceder aos mesmos dados.

- um Estado-Membro poderá ter aprovado legislação que exija a publicação dos nomes das entidades que tenham feito doações a partidos políticos de montante superior a um determinado limite. As informações suscetíveis de revelar as convicções políticas dos doadores são publicadas num sítio *Web* criado para o efeito. São aplicáveis garantias técnicas, tais como limites ao descarregamento de grandes quantidade de dados e às capacidades de pesquisa. Um grupo de ativistas solicita acesso a uma grande quantidade de dados para reutilização ao abrigo da Diretiva ISP, com vista a criar um novo sítio *Web* com funcionalidades adicionais e melhores capacidades de pesquisa.
- o nome e a morada do proprietário de um imóvel são públicos no registo predial de um Estado-Membro, mas a navegação na base de dados de acesso público é limitada, sendo apenas permitida a pesquisa por imóvel e não por proprietário. Também existem limites ao descarregamento de grandes quantidades de dados. Uma empresa solicita acesso a grandes quantidades de dados para reutilização, com vista a criar um sítio *Web* de utilização mais fácil, a um preço mais competitivo.
- os registos comerciais de um Estado-Membro permitem o acesso do público a uma grande variedade de dados pessoais, incluindo nomes, moradas e assinaturas dos administradores, bem como a informações sobre os proprietários de certos tipos de empresas. Existem algumas restrições às capacidades de pesquisa e limites ao número de dados que podem ser descarregados. As informações estão disponíveis através de um sítio *Web* criado para o efeito, mediante o pagamento de uma taxa. Uma empresa solicita acesso a grandes quantidades de dados para reutilização, tendo em vista a criação de um sítio *Web* que combina informações de vários tipos de registos e a disponibilização de melhores informações a um preço mais competitivo.

Em todos estes casos, o organismo do setor público em causa tem de realizar uma cuidadosa avaliação do impacto na proteção de dados para decidir se os dados podem ser disponibilizados para reutilização ao abrigo da Diretiva ISP e, em caso afirmativo, determinar se a legislação sobre proteção de dados impõe a adoção de condições e garantias específicas. O «princípio da reutilização» não é de aplicação automática e não prevalece sobre as disposições aplicáveis da legislação sobre proteção de dados.

Esta cuidadosa avaliação é particularmente importante porque, nos termos da Diretiva ISP, o organismo do setor público, em princípio, não pode ter em consideração a identidade do reutilizador que solicita o acesso. Nos termos do artigo 10.º (Não discriminação), «as eventuais condições aplicáveis à reutilização de documentos não devem ser discriminatórias para categorias de reutilização equivalentes». Além disso, nos termos do artigo 11.º (Proibição de acordos exclusivos), «a reutilização de documentos está aberta a todos os potenciais intervenientes no mercado... Os contratos ou outros acordos celebrados entre organismos do setor público que possuam esses documentos e terceiros não criam direitos exclusivos.»

Por conseguinte, ao tomarem a decisão de autorizar ou não a reutilização, os organismos do setor público têm de ponderar a compatibilidade da autorização de reutilização ao abrigo de uma licença aberta, não apenas em relação ao requerente como também a qualquer pessoa que solicite acesso aos dados. Isso exige um elevado grau de convicção de que nenhum dos potenciais reutilizadores poderá utilizar abusivamente os dados pessoais disponibilizados.

A Diretiva ISP não exclui a possibilidade de serem estipulados termos e condições que apenas autorizem o tratamento para fins específicos. O organismo do setor público depara-se então com a questão da compatibilidade da reutilização por qualquer «potencial interveniente no mercado» para estes fins com os fins especificados pelo referido organismo. A potencial reutilização de informações sobre o pagamento de impostos por instituições financeiras, por exemplo, para fins de elaboração de relatórios de crédito, é relevante, dado que, de acordo com o critério «qualquer pessoa», aquelas constituem um potencial reutilizador. Consequentemente, a fim de responder às preocupações em matéria de proteção de dados e, em especial, para assegurar o respeito pelo princípio da limitação da finalidade, o organismo do setor público (ou o legislador) deve poder estabelecer limites às finalidades da reutilização, quando necessário.

VI. Reutilização de conjuntos de dados agregados e anonimizados obtidos a partir de dados pessoais

6.1. Quais são os benefícios da agregação e anonimização para a reutilização de ISP?

Até à data, as iniciativas de reutilização de ISP lançadas por organismos do setor público através de «portais de dados abertos» ou de outras plataformas visam, por norma, a disponibilização de dados agregados e anonimizados para reutilização, e não dados pessoais enquanto tais. Esta abordagem é, de facto, mais segura e deve ser incentivada.

Geralmente, a legislação sobre proteção de dados não permite que os organismos do setor público divulguem publicamente dados pessoais que foram recolhidos para outro fim, normalmente de caráter administrativo[22]. Assim, nestes casos, a sua reutilização no âmbito das iniciativas de reutilização de ISP também não é possível. Não são geralmente dados pessoais, mas sim dados estatísticos obtidos a partir de dados pessoais que são e que devem, em princípio, ser disponibilizados para reutilização. Esta é a solução mais eficaz para minimizar os riscos de uma divulgação acidental de dados pessoais. Estes conjuntos de dados anonimizados e agregados não devem permitir a reidentificação das pessoas e, como tal, não devem conter dados pessoais.

Decidir qual o nível de agregação adequado e quais a técnicas de anonimização específicas a utilizar não é uma tarefa fácil. Se a agregação e a anonimização não forem realizadas eficazmente, existe o risco de as pessoas serem, ainda assim, reidentificadas a partir desses conjuntos de dados. Por conseguinte, a legislação sobre proteção de dados tem um importante papel a desempenhar, dado que pode ajudar a fixar o nível de segurança exigido quando está em causa a divulgação de dados anonimizados e agregados no âmbito de uma iniciativa ISP.

A Diretiva 95/46/CE estabelece um elevado grau de anonimização

Para efeitos do presente documento, o termo «anonimização» refere-se a dados que já não podem ser considerados dados pessoais na aceção do artigo 2.º, alínea a), da Diretiva 95/46/CE. Nos termos desta disposição, entende-se por «dados pessoais», «qualquer informação relativa a uma pessoa singular identificada ou identificável («titular dos dados»); é considerado identificável todo aquele que possa ser identificado, direta ou indiretamente, nomeadamente por referência a um número de

[22] É evidente que, em certos casos, a legislação sobre liberdade de informação poderá exigir a divulgação de dados pessoais, e o interesse na transparência e a disponibilidade de informações em algumas situações poderão sobrepor-se às preocupações em matéria de privacidade e proteção de dados. Trata-se de uma área em evolução, que poderá trazer alterações no futuro.

identificação ou a um ou mais elementos específicos da sua identidade física, fisiológica, psíquica, económica, cultural ou social»[23].

O considerando 26 da Diretiva 95/46/CE também é relevante, estabelecendo ainda que, «para determinar se uma pessoa é identificável, importa considerar o conjunto dos meios suscetíveis de serem razoavelmente utilizados, seja pelo responsável pelo tratamento, seja por qualquer outra pessoa, para identificar a referida pessoa».

Deve salientar-se que, deste modo, é exigido um elevado grau de anonimização, tal como discutido mais adiante no presente parecer. A menos que seja possível atingir este grau de anonimização dos dados, a legislação sobre proteção de dados continua a ser aplicável. Tal significa, por exemplo, que, neste caso, a divulgação pública das informações (e qualquer utilização posterior) tem de ser «compatível» com as finalidades iniciais da recolha dos dados ao abrigo do artigo 6.º, n.º 1, alínea b), da Diretiva 95/46/CE. Além disso, tem de existir uma base legal adequada para o tratamento nos termos do artigo 7.º, alíneas a) a f), desta Diretiva (por exemplo, consentimento ou necessidade para cumprir a lei). Em contrapartida, se os dados tiverem sido anonimizados na aceção do artigo 2.º, alínea a), e do considerando 26 da Diretiva 95/46/CE, as regras sobre proteção de dados já não serão aplicáveis e os reutilizadores poderão reutilizar os dados sem estarem sujeitos a estas restrições.

Mais uma vez, é de sublinhar que, para efeitos do presente parecer, o termo «dados anonimizados» designa dados que já não são considerados dados pessoais. Em especial, importa distinguir dados anonimizados de dados que foram manipulados com recurso a várias técnicas para minimizar os riscos de reidentificação das pessoas em causa, mas que não atingiram o grau de anonimização exigido pelo artigo 2.º, alínea a), e pelo considerando 26 da Diretiva 95/46/CE[24]. Em muitos cenários, o recurso a estas técnicas só é adequado em casos de divulgação limitada para reutilização por terceiros que foram submetidos a um processo de controlo prévio, mas não para divulgação pública total e reutilização ao abrigo de uma licença aberta.

É igualmente importante salientar que, assim que os dados são divulgados ao público para reutilização, deixa de ser possível controlar quem pode aceder a esses dados. A probabilidade de «qualquer outra pessoa» possuir e utilizar os meios para reidentificar as pessoas em causa aumentará muito significativamente. Por conseguinte, e independentemente da interpretação dada ao considerando 26 noutros contextos, estando em causa a disponibilização de dados para reutilização ao abrigo da Diretiva ISP, o GT29 deseja deixar bem claro que devem ser tomadas todas as precauções possíveis para assegurar que os conjuntos de dados que serão divulgados não contêm dados suscetíveis de serem reidentificados através de meios que, com uma probabilidade razoável, possam ser utilizados por qualquer pessoa, incluindo potenciais reutilizadores, mas também outras entidades com interesse em obter os dados, nomeadamente as autoridades policiais.

[23] Na sua declaração de 27 de fevereiro de 2012 relativa ao atual debate sobre o pacote da reforma da proteção de dados, o GT29 salientou que uma pessoa singular pode ser considerada identificável quando, num grupo de pessoas, possa ser distinguida das outras e, consequentemente, tratada de forma diferente. Assim, o conceito de identificabilidade abrange a distinção. A declaração esclarece igualmente que os números de identificação, os dados sobre a localização, os endereços IP, os identificadores em linha e outros fatores específicos relacionados com uma pessoa devem ser considerados dados pessoais.

[24] A declaração de 27 de fevereiro de 2013 salienta que, sempre que for possível estabelecer a identidade de uma pessoa por regressão ou identificá-la (indiretamente) por outros meios, as regras sobre proteção de dados continuam a ser aplicáveis.

Orientações adicionais sobre a anonimização e o conceito de dados pessoais

Para mais orientações sobre a anonimização e o conceito de dados pessoais, ver o Parecer 4/2007 do GT29 sobre o conceito de dados pessoais, adotado em 20 de junho de 2007 (WP 136). O GT29 poderá fornecer orientações adicionais sobre técnicas de anonimização num outro documento, no segundo semestre de 2013.

6.2. Quais são os desafios e os limites da anonimização para a reutilização de ISP?

Com a evolução da tecnologia informática moderna e a disponibilidade ubíqua de informação, a anonimização torna-se cada vez mais difícil. A reidentificação das pessoas é uma ameaça cada vez mais comum e presente.[25] Na prática, existe uma área cinzenta muito significativa, em que o responsável pelo tratamento que divulga os dados poderá pensar que o conjunto de dados está anonimizado, mas um terceiro poderá, ainda assim, conseguir identificar algumas pessoas a partir desses dados, utilizando, por exemplo, outras informações publicamente disponíveis ou outras informações que estejam ao seu dispor.

Um dos maiores fatores de risco é o crescente volume de dados em linha e fora de linha, quer publicamente disponíveis, quer concentrados nas mãos de organizações comerciais, que podem depois ser utilizados para criar perfis individuais para publicidade comportamental e para um leque cada vez mais diversificado de outros fins. Se forem combinadas com o grande volume de dados («*big data*») que estas organizações têm já à sua disposição, as ISP obtidas a partir de dados pessoais e disponibilizadas para reutilização podem aumentar a probabilidade de os indivíduos serem agora identificados ou de os seus perfis serem alargados, muitas vezes sem que estes se apercebam do que está a acontecer.

6.3. Quem deve realizar a agregação e anonimização e quando?

A agregação e a anonimização devem ter lugar tão cedo quanto possível e ser realizadas pelo responsável pelo tratamento ou por um terceiro de confiança que atue em nome de um ou de vários responsáveis pelo tratamento (e que também possua as competências especializadas necessárias). A anonimização não pode ser deixada a cargo do reutilizador (por exemplo, como condição da licença). Além disso, é importante garantir que o possível terceiro que realizará a agregação e anonimização não seja afetado por qualquer conflito de interesses e seja claramente responsável por assegurar que os dados pessoais só serão utilizados para realizar a anonimização e que serão aplicadas todas as garantias que se mostrarem necessárias para este efeito. O terceiro deve ainda estar em condições de garantir que os dados pessoais a partir dos quais foram obtidos os conjuntos de dados agregados e anonimizados serão eliminados logo que deixem de ser necessários para aquele fim.

---

[25] Ver, por exemplo, um relatório elaborado em 2011 para o Gabinete do Primeiro-Ministro do Reino Unido por Kieron O'Hara da Universidade de Southampton, intitulado «*Transparent Government, Not transparent Citizens*» (Um Governo Transparente e não Cidadãos Transparentes), onde o autor alerta para a possibilidade de identificar indivíduos a partir de dados anonimizados, por exemplo, juntando diferentes informações provenientes de diferentes fontes como se tratassem das peças de um *puzzle* (a chamada «*jigsaw identification*») e afirma que não existem soluções técnicas completas para o problema da «desanonimização». Disponível em: http://www.cabinetoffice.gov.uk/sites/default/files/resources/transparency-and-privacy-review-annex-b.pdf. Ver também «*Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*» (Promessas de privacidade quebradas: a resposta ao surpreendente fracasso da anonimização), de Paul Ohm da Faculdade de Direito da Universidade do Colorado, 57 UCLA Law Review 1701 (2010), disponível em linha em: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

6.4. Avaliar os riscos de reidentificação

A menos que seja possível anonimizar os dados na aceção do artigo 2.º, alínea a), e do considerando 26 da Diretiva 95/46/CE, a legislação sobre proteção de dados continuará a ser aplicável.

Os responsáveis pelo tratamento devem determinar se é possível, em termos razoáveis, identificar um indivíduo a partir do conjunto de dados «anonimizado» que será disponibilizado e a partir de outros dados ou, por outras palavras, se uma organização ou indivíduo poderia identificar um indivíduo a partir dos dados a divulgar – isoladamente ou em combinação com outras informações disponíveis.

Tal como explicado na secção 6.1, o presente parecer não pretende fornecer orientações exaustivas e conclusivas sobre a avaliação dos riscos de reidentificação. Do mesmo modo, também não visa fornecer uma definição conclusiva dos termos «anonimização» ou «dados anonimizados». No entanto, relembra que o leitor poderá encontrar orientações adicionais noutros documentos (incluindo os documentos indicados na secção 6.1) e que, tal como referido na secção 6.1 e na secção 2.2, o subgrupo «Tecnologia» do GT29 está atualmente a desenvolver um trabalho sobre técnicas de anonimização.

Dito isto, e sem pretensões de exaustividade, o GT29 gostaria de salientar alguns dos fatores/conceitos que seria útil ponderar durante a avaliação dos riscos de reidentificação, incluindo, em especial:

- outros dados que estejam à disposição do público em geral ou de outros indivíduos ou organizações e a suscetibilidade de os dados a publicar poderem ser associados a outros conjuntos de dados;
- a probabilidade de serem efetuadas tentativas de reidentificação (alguns tipos de dados despertarão mais o interesse de potenciais intrusos do que outros); e
- a probabilidade de sucesso de eventuais tentativas de reidentificação, tendo em conta a eficácia das técnicas de anonimização propostas[26].

Que «outras» informações estão disponíveis?

Para determinar se um indivíduo pode ser indiretamente identificado, é necessário considerar se é possível proceder a essa identificação utilizando os dados em questão (no nosso caso, o conjunto de dados «anonimizado») ou partindo desses dados e de *outras informações* que estejam (ou seja possível ou provável que venham a estar) na posse da organização ou indivíduo que tenta proceder à identificação.

As «outras informações» necessárias para a reidentificação poderão ser informações à disposição de certas empresas ou outras organizações, incluindo autoridades policiais ou outros organismos do setor público, de certos indivíduos ou do público em geral, caso tenham sido publicadas na Internet, por exemplo. Um exemplo óbvio é a combinação de dados publicamente disponíveis (como os cadernos eleitorais, a lista telefónica ou outros dados que podem ser obtidos mediante uma pesquisa na Internet) com os dados (incorretamente) «anonimizados», permitindo a identificação de um indivíduo (por exemplo, utilizando a sua data de nascimento e o seu código postal).

[26] Para mais informações sobre técnicas de anonimização, ver o parecer do GT29 sobre esta matéria a publicar brevemente.

Os riscos de reidentificação podem aumentar quando um indivíduo ou grupo de indivíduos já dispõe de muitas informações sobre outro indivíduo, por exemplo um familiar, um colega, um contacto numa rede social, um médico, um professor, um agente da autoridade ou outro profissional.

Porém, o que importa aqui não é simplesmente se o indivíduo com conhecimento prévio pode identificar a pessoa em causa, mas sim se ficará a saber algo de novo a partir das informações obtidas através de reidentificação. Os dois exemplos que se seguem ilustram a importância desta distinção.

Exemplo um: estatísticas sobre sarampo. Num caso, dados estatísticos anonimizados poderão revelar que, na localidade A, no ano de 2012, X pessoas contraíram sarampo. Os dados não são desagregados, nem são fornecidas mais informações. Um médico que contribuiu para as estatísticas fornecendo informações sobre os seus próprios pacientes às autoridades de saúde competentes possui dados mais completos sobre estes pacientes no seu consultório, que estão protegidos pelo sigilo médico. O médico poderia reidentificar facilmente vários pacientes a partir do conjunto de dados estatísticos. Da mesma forma, uma mãe cujo filho tivesse contraído sarampo nesse ano poderia facilmente reidentificá-lo nesse conjunto de dados. No entanto, nem a mãe nem o médico ficariam a saber algo que não sabiam antes de o conjunto de dados anonimizados ter sido disponibilizado ao público.

Exemplo dois: abuso de drogas e álcool, abuso sexual e desempenho escolar. O exemplo anterior distingue-se da seguinte situação. É realizado um estudo sobre a correlação entre o abuso de drogas e álcool pelos progenitores, o abuso sexual de crianças e o desempenho escolar. Na sequência deste estudo, são publicados dados alegadamente «anonimizados» com boas intenções, mas sem uma avaliação cuidadosa dos riscos de reidentificação.

As estatísticas revelam, designadamente, que na Escola A, onde estão inscritos 500 alunos, no ano de 2012, 20 % dos alunos (100 alunos) viviam em agregados familiares em que, pelo menos, um dos progenitores era alcoólico ou toxicodependente. Destes, em 8 % dos casos (8 alunos), a criança tinha sido vítima de abuso sexual. O relatório especifica ainda que mais nenhum aluno tinha sido vítima de abuso sexual na Escola A.

Os números mostram também que, em 96 % dos casos (96 alunos), as crianças cujos progenitores eram alcoólicos ou toxicodependentes tinham sérios problemas ao nível do desempenho escolar (sendo considerados «maus alunos» de acordo com um critério académico adequado); no entanto, nesta escola, apenas 50 % das crianças vítimas de abuso sexual (4 alunos) tinham dificuldades significativas com o trabalho escolar.

Na escola, é do conhecimento geral que AA, um aluno inteligente e trabalhador, vem de uma família problemática e que a sua mãe é alcoólica. É frequentemente vítima de intimidação pelos seus colegas. Ao lerem as estatísticas republicadas no jornal da escola, estes mesmos colegas apercebem-se agora de que AA deve pertencer ao grupo de 50 % de crianças vítimas de abuso sexual que não têm dificuldades na escola («bons alunos»). Deste modo, adquiriram novas informações (e, neste caso, muito sensíveis) a partir de um conjunto de dados deficientemente anonimizado.

O risco de combinar informações para obter dados pessoais aumenta com o desenvolvimento das técnicas de ligação de dados e da capacidade computacional e com a disponibilização ao público de um número cada vez maior de informações potencialmente «combináveis». Com efeito, a capacidade computacional duplica todos os anos e o armazenamento de dados, devido também à existência de serviços em nuvem, tornar-se-á provavelmente uma mercadoria transacionável. Por conseguinte, o risco de reidentificação através da ligação de dados é imprevisível porque nunca será

possível determinar, com segurança, que dados estão já disponíveis e que dados poderão ser divulgados no futuro.

Não obstante toda esta incerteza, os riscos de reidentificação podem ser geralmente ser minimizados, pelo menos em parte, mediante o respeito pelo princípio da minimização dos dados, ou seja, assegurando que apenas são divulgados os dados necessários para um determinado fim.

A probabilidade de sucesso das tentativas de reidentificação: o critério do «intruso motivado»

O critério do «intruso motivado» é um conceito novo, cuja validade ainda não está plenamente comprovada. Poderá ser útil para determinar:

- se alguém teria a motivação para realizar a reidentificação; e
- a possibilidade/probabilidade de sucesso da reidentificação.

O critério do intruso motivado passa essencialmente por determinar se um «intruso» conseguiria proceder à reidentificação *se* estivesse motivado para o tentar. O «intruso motivado» é uma pessoa (singular ou coletiva) que deseja identificar o indivíduo cujos dados pessoais serviram de base aos dados anonimizados. Este critério visa determinar se o intruso motivado seria bem-sucedido. Nesta abordagem, presume-se que o «intruso motivado» é competente e tem acesso a recursos proporcionais à sua motivação para a reidentificação.

Alguns tipos de dados terão mais interesse para um «intruso motivado» do que outros. Por exemplo, um intruso, em geral, poderá estar mais motivado para reidentificar dados pessoais que:

- tenham um valor comercial considerável (nomeadamente no mercado negro ou fora da União Europeia) e, como tal, possam ser comprados e vendidos com fins lucrativos[27];
- possam ser utilizados pelas autoridades policiais ou pelos serviços de informação;
- revelem informações com interesse mediático sobre figuras públicas;
- possam ser utilizados para fins políticos ou ativistas (por exemplo, como parte de uma campanha contra uma determinada organização ou pessoa);
- poderiam ser utilizados por motivos pessoais condenáveis (por exemplo, perseguição obsessiva, assédio, intimidação ou apenas para humilhar terceiros);
- poderiam despertar a curiosidade (por exemplo, o desejo de um habitante local de descobrir quem esteve envolvido num incidente apresentado num mapa da criminalidade).

Embora seja útil considerar as potenciais motivações dos potenciais intrusos, o GT29 salienta que esta abordagem também tem limites consideráveis:

- o exercício poderá ser, em certa medida, especulativo.
- na ausência de «fatores de motivação» óbvios, tais como os descritos em cima, o exercício poderá criar uma falsa sensação de segurança e sugerir que é possível disponibilizar para reutilização dados pessoais relativamente inócuos sem uma anonimização eficaz.

[27] Poderão incluir, por exemplo, dados transacionais ou outros dados comportamentais dos quais seja possível inferir perfis de consumo individuais, que poderão depois ser utilizados para fins publicitários ou discriminação de preços; informações financeiras ou outras informações que viabilizem o furto de identidade; informações sensíveis que poderão ser utilizadas para efeitos de chantagem ou de discriminação; informações médicas suscetíveis de serem utilizadas pelas companhias de seguros, por exemplo, para recusar a cobertura com fundamento numa doença preexistente; informações que possibilitem inferências sobre solvabilidade que poderão ser utilizadas para avaliar riscos de crédito; etc.

- podem existir intrusos sofisticados, inovadores e pioneiros, que encontrem utilizações para dados desidentificados que não sejam óbvias para outras pessoas.
- com a crescente tendência no sentido da analítica de grandes volumes de dados («*big data*»), existe um risco cada vez maior de dados aparentemente inócuos, uma vez desidentificados e combinados com outras informações, acabarem por colocar riscos mais graves.

6.5. Teste da reidentificação

Em alguns casos, pode ser difícil estabelecer o risco de identificação, especialmente quando um terceiro utiliza métodos estatísticos complexos para combinar vários dados anonimizados. Consequentemente, durante a avaliação global realizada para identificar o risco de reidentificação, é boa prática realizar um teste de reidentificação – um tipo de teste de «penetração» – para detetar e gerir vulnerabilidades de reidentificação. Este teste consiste em tentar reidentificar indivíduos a partir dos conjuntos de dados que se pretende divulgar.

A primeira fase do processo de teste da reidentificação deve consistir num inventário dos conjuntos de dados que o organismo do setor público publicou ou tenciona publicar. Seguidamente, importa tentar determinar que outros dados – pessoais ou não – estão disponíveis e cuja ligação aos dados em causa poderia resultar em reidentificação. Em especial, a realização de «testes de penetração» com fins específicos poderá ajudar a avaliar os riscos de identificação através da colagem de diferentes informações para criar uma imagem mais completa de alguém.

Naturalmente, o teste da reidentificação não deve ser considerado uma panaceia e não deve conduzir a uma falsa sensação de segurança. Em primeiro lugar, poderá ser difícil realizar o teste, dado que exige frequentemente conhecimentos técnicos significativos e ferramentas adequadas, bem como uma noção dos dados que poderão estar disponíveis. Em segundo lugar, os responsáveis pelo tratamento dos dados também devem estar cientes de que o risco de identificação pode mudar ao longo do tempo. Por exemplo, estão agora disponíveis ferramentas e técnicas de análise de dados cada vez mais sofisticadas e a preços acessíveis e a correlação com outros conjuntos de dados torna-se cada vez mais fácil à medida que são gerados cada vez mais dados. Por este motivo, as organizações devem rever periodicamente a sua política sobre a divulgação de dados e as técnicas de anonimização utilizadas. Além disso, as decisões nunca devem ter exclusivamente por base as ameaças existentes, mas também as ameaças futuras previsíveis.

Uma vez realizada a avaliação dos riscos de reidentificação descrita na secção 6.4 e, quando necessário, depois de realizado o teste da reidentificação, o organismo do setor público está em condições de decidir se o conjunto de dados pode ser considerado anonimizado, ou seja, se já não contém dados pessoais na aceção do artigo 2.º, alínea a) e do considerando 26 da Diretiva 95/46/CE. Em caso afirmativo, o conjunto de dados pode ser divulgado sem estar sujeito a restrições relacionadas com a proteção de dados[28]. Por outro lado, se o teste da reidentificação for bem-sucedido, estes dados não poderão (ou deixarão de poder) ser disponibilizados como dados anonimizados, tendo então de ser considerados dados pessoais (e, assim sendo, a sua divulgação poderá não ser possível ou poderá apenas ser possível nas condições referidas na secção VII).

[28] Ver, no entanto, a secção 10.3 sobre «Condições da licença para conjuntos de dados anonimizados» e, em especial, a necessidade de impor garantias para continuar a assegurar que os indivíduos não serão reidentificados.

6.6. Retirada de conjuntos de dados comprometidos

Na eventualidade de reidentificação comprovada de dados a partir de um conjunto de dados aberto, o organismo do setor público que forneceu o conjunto de dados tem de poder desativar a alimentação dos dados (*feed*) ou retirar o conjunto de dados do sítio *Web* de dados abertos. Se retirar o conjunto de dados do sítio *Web*, o organismo do setor público também tem de informar os reutilizadores e contactá-los no sentido de interromperem o tratamento e eliminarem todos os dados provenientes do conjunto de dados comprometido. Uma vez que será difícil informar todos os reutilizadores ao abrigo do regime de licenciamento aberto exigido pela Diretiva ISP, os organismos públicos têm de adotar medidas razoavelmente eficazes para responder a esta questão. Embora a retirada possa muitas vezes não ir a tempo de evitar os danos, é uma medida necessária para ajudar a minimizar eventuais efeitos adversos sobre as pessoas em causa.

VII. **Disponibilização de dados pessoais para reutilização**

7.1. Exemplos de dados pessoais publicamente disponíveis publicados por organismos do setor público

Embora as iniciativas de reutilização de ISP assumam habitualmente a forma de disponibilização de conjuntos de dados anonimizados, em alguns casos os organismos do setor público também poderão disponibilizar dados pessoais para reutilização.

Muitos registos publicamente disponíveis, como os registos prediais e os registos comerciais, contêm grandes quantidades de dados pessoais e, devido às iniciativas de administração em linha, estão também cada vez mais disponíveis na Internet. Existem muitos outros exemplos de casos em que os legisladores, em especial os Estados-Membros, estabeleceram uma base legal para disponibilizar dados pessoais na Internet ou mediante um pedido de acesso a documentos. Estes incluem, por exemplo[29]:

- despesas, salários ou declarações de conflito de interesses de certos titulares de cargos públicos ou de beneficiários de auxílios estatais (por exemplo, subsídios agrícolas);
- nomes das organizações ou indivíduos que fazem doações a partidos políticos;
- declarações fiscais de pessoas singulares[30];
- decisões judiciais (sendo os nomes das partes ou de outros indivíduos por vezes suprimidos ou substituídos por iniciais para reduzir o risco de reidentificação);
- listas eleitorais;
- listas das audiências judiciais que terão lugar em determinados dias.

Em cada um destes casos, os organismos do setor público ou os legisladores poderão considerar proativamente se pretendem disponibilizar estes dados para reutilização (por exemplo, com vista a melhorar serviços públicos como o acesso a registos comerciais ou prediais). Os organismos do setor público também poderão ser contactados por potenciais reutilizadores que solicitem a reutilização dos dados. Noutros casos, também é possível que os potenciais reutilizadores acedam simplesmente aos dados pessoais que já estão disponíveis em linha e os utilizem sem contactarem necessariamente o organismo do setor público que publicou as informações. Nestes três casos, os reutilizadores teriam obviamente de cumprir a legislação sobre proteção de dados, pois estariam em causa dados pessoais.

---

[29] Ver também os exemplos apresentados na secção V, a propósito do âmbito de aplicação da Diretiva ISP.

[30] Ver, por exemplo, o acórdão do Tribunal de Justiça da União Europeia, de 16 de dezembro de 2008, no processo C-73/07, *Tietosuojavaltuutettu/Satakunnan Markkinapörssi Oy en Satamedia Oy*.

7.2. Diferenças nos regimes nacionais de acesso

As obrigações legais de disponibilizar ao público certos dados pessoais variam significativamente entre os Estados-Membros devido a diferenças nas tradições jurídicas e culturais. Alguns Estados-Membros estabeleceram uma base legal para disponibilizar certos dados pessoais, enquanto outros proíbem a divulgação destes mesmos dados pessoais na mesma situação. A Diretiva ISP reconhece e deixa bem claro que tem por base os regimes de acesso em vigor nos Estados-Membros e que não altera as regras nacionais de acesso a documentos.[31]

7.3. Necessidade de uma avaliação do impacto na proteção de dados e de garantias adequadas

Em regra, sempre que seja ponderada a disponibilização de dados pessoais para reutilização, é indispensável adotar uma abordagem cautelosa. O GT29 recomenda, em especial, a realização de uma minuciosa avaliação do impacto na proteção de dados antes da publicação do conjunto de dados (ou antes da adoção de uma lei que exija a publicação), que analise igualmente as possibilidades e o potencial impacto da reutilização. De um modo geral, deve ser evitada a disponibilização de dados pessoais para reutilização ao abrigo de uma licença aberta sem quaisquer restrições técnicas e jurídicas.

7.4. Importância de um regime de licenciamento

O GT29 recomenda ainda a implementação de um rigoroso regime de licenciamento, que tem de ser corretamente aplicado para evitar a utilização de dados pessoais para fins incompatíveis – por exemplo, para mensagem comerciais não solicitadas ou de forma que as pessoas em causa considerassem inesperada, imprópria ou criticável.

7.5. Importância de uma base legal sólida para publicação e também para reutilização

O GT29 reitera a importância de estabelecer uma base legal sólida para disponibilizar ao público dados pessoais, tomando em consideração as regras sobre proteção de dados aplicáveis, nomeadamente o princípio da proporcionalidade, da minimização dos dados e da limitação da finalidade.

O GT29 recomenda que qualquer legislação que implique o acesso público a dados especifique claramente os fins para que poderão ser divulgados dados pessoais. Se estes fins não forem especificados, ou se o forem em termos vagos e gerais, a previsibilidade e a segurança jurídica ficam comprometidas. Em especial, perante um pedido de reutilização, será muito difícil para o organismo do setor público e para os potenciais reutilizadores em causa determinar quais eram as finalidades iniciais pretendidas com a publicação e, subsequentemente, que outras finalidades seriam compatíveis com essas finalidades iniciais. Tal como já mencionado, mesmo que os dados pessoais tenham sido publicados na Internet, não se deve presumir que podem ser objeto de um tratamento posterior para qualquer fim possível.

Nestes casos, qualquer reutilização posterior deverá ter uma base legal adequada (por exemplo, consentimento ou exigência legal) nos termos do artigo 7.º, alíneas a) a f), da Diretiva 95/46/CE e cumprir todos os outros princípios da proteção de dados.

[31] Não obstante, conforme explicado na secção 5.4, a legislação nacional tem de cumprir, em qualquer caso, o artigo 8.º da CEDH e os artigos 7.º e 8.º da Carta da UE, tal como interpretados pela jurisprudência relevante.

7.6. Limitação da finalidade

Não é fácil aplicar eficazmente o princípio da limitação da finalidade no caso de reutilização de ISP. Por um lado, a própria ideia e força motriz da inovação subjacente ao conceito de «dados abertos» e da reutilização de ISP é a de disponibilizar informações para reutilização tendo em vista produtos e serviços novos e inovadores e, como tal, para fins que não estão previamente definidos e não podem ser previstos com exatidão. A Diretiva ISP também exige que as licenças não coloquem restrições desnecessárias à reutilização.

Por outro lado, a limitação da finalidade é um dos princípios fundamentais da proteção de dados, segundo o qual os dados pessoais recolhidos para uma determinada finalidade não devem ser posteriormente utilizados para outra finalidade, incompatível com a primeira.[32] Este princípio é igualmente aplicável aos dados pessoais publicamente disponíveis. O simples facto de esses dados pessoais estarem publicamente disponíveis para um fim específico não significa que estejam disponíveis para reutilização para qualquer outro fim.

Por exemplo, as despesas dos titulares de altos cargos públicos são disponibilizadas na Internet por razões de transparência, mas permitir a sua reutilização por qualquer membro do público para outros fins poderá não ser compatível.

Tal como discutido em maior pormenor no Parecer 3/2013 do GT29 sobre a limitação da finalidade (ver secção III.2.2 e anexo 1), a determinação da compatibilidade do tratamento posterior de dados pessoais com as finalidades para as quais esses dados foram recolhidos exige a avaliação de vários fatores. Importa ter em conta, em especial:

a) A relação entre as finalidades para as quais os dados pessoais foram recolhidos e as finalidades do tratamento posterior;
b) O contexto em que os dados pessoais foram recolhidos e as expectativas razoáveis das pessoas em causa quanto à sua utilização posterior;
c) A natureza dos dados pessoais e o impacto do tratamento posterior sobre as pessoas em causa;
d) As garantias aplicadas pelo responsável pelo tratamento para assegurar um tratamento leal e evitar um impacto indevido sobre as pessoas em causa.

É necessário analisar estes fatores-chave quando é tomada a decisão de disponibilizar ou não publicamente quaisquer dados pessoais, bem como, em cada caso, quando estiver em causa a reutilização de dados pessoais. Eis alguns exemplos:

- um organismo do setor público publica um diretório com as informações de contacto dos seus funcionários, incluindo nome, cargo, morada do local de trabalho e número de telefone do local de trabalho. A finalidade óbvia – embora não indicada expressamente – deste diretório é ajudar o público a identificar a pessoa que devem contactar para pedidos de informações e outros assuntos de caráter oficial. Um reutilizador pretende «colher» o conteúdo deste diretório, combiná-lo com as moradas e os números de telefone de casa dos funcionários (quando estejam publicamente disponíveis, por exemplo, numa lista telefónica), e disponibilizar tanto a morada e o número de telefone do trabalho como de casa num mapa interativo para mostrar onde vivem e trabalham diferentes funcionários públicos. Esta combinação e reutilização de dados tem de ser considerada incompatível com a finalidade

[32] Esses dados só podem ser utilizados de forma incompatível com as finalidades especificadas no momento da recolha em casos excecionais e sob reserva de serem impostas de rigorosas garantias nos termos do artigo 13.º da Diretiva 95/46/CE. Ver secção III.3 do Parecer 3/2013 do GT29 sobre a limitação da finalidade.

inicial. Um funcionário público cujas informações de contacto profissional são divulgadas para que possa ser contactado pelo público não esperaria, em termos razoáveis, que estas informações fossem posteriormente correlacionadas com outros dados que tivesse disponibilizado publicamente para outra finalidade não relacionada com o seu trabalho.

- em alguns Estados-Membros, a legislação nacional exige a publicação de editais de casamento que podem ser consultados por qualquer pessoa. Essa publicação tem por objetivo divulgar a vontade dos nubentes de contrair casamento e dar aos interessados a oportunidade de manifestarem a sua oposição. O facto de os dados pessoais contidos nos editais de casamento poderem ser consultados por qualquer pessoa não autoriza terceiros a utilizarem essas informações para enviarem mensagens comerciais aos casais. Esta utilização adicional seria incompatível à luz do objetivo da publicação destes editais, que consiste em apresentar objeções ao casamento nos termos da lei.

7.7. Fins comerciais *versus* fins não comerciais

O Parecer 7/2003 destaca as atividades comerciais como o principal incentivo para a reutilização de ISP, por oposição ao acesso à informação, em que o objetivo da legislação sobre liberdade de informação consiste em garantir transparência, abertura e responsabilização perante os cidadãos.

O Parecer 7/2003 salienta ainda que «regra geral, a informação é utilizada para fins privados e não comerciais». Esta afirmação deve ser atualizada à luz da experiência entretanto adquirida com a reutilização de ISP. A experiência com iniciativas de dados abertos revela que a reutilização de ISP poderá também contribuir significativamente para reforçar a transparência e a responsabilização, podendo igualmente resultar numa melhor utilização dos serviços públicos. A distinção entre reutilização para fins comerciais e não comerciais não deve ser um critério decisivo na análise da compatibilidade de utilizações posteriores de dados pessoais. Esta análise não deve assentar primordialmente no facto de o modelo económico de um potencial reutilizador ter ou não fins lucrativos.

O que importa analisar cuidadosamente é a compatibilidade das finalidades e da forma como os dados são tratados posteriormente com as finalidades iniciais de acordo com os critérios mencionados na secção 7.6. No caso de reutilização de ISP, esta análise conduz inevitavelmente à ponderação de uma série de cenários de tratamento, e não apenas de um.

7.8. Proporcionalidade e outras considerações

Outro dos princípios fundamentais previstos na Diretiva 95/46/CE é o da proporcionalidade[33]. Existem muitos métodos e modalidades diferentes de disponibilizar dados pessoais ao público. Alguns deles são mais intrusivos do que outros e apresentam mais riscos. Consequentemente, alguns poderão ser considerados proporcionados, outros não.

Tal como acontece com a finalidade, o controlo do tratamento posterior dos dados e a garantia do cumprimento de outros princípios consagrados na legislação sobre proteção de dados (como, por exemplo, a proporcionalidade) são questões que suscitam preocupação. Depois de os dados terem sido disponibilizados ao público, especialmente na Internet, é muito difícil limitar eficazmente a sua utilização e garantir o cumprimento da legislação sobre proteção de dados.

Alguns dos desafios que se colocam em termos de garantia do cumprimento da legislação sobre proteção de dados incluem:

[33] Ver artigo 6.º, n.º 1, alínea c), da Diretiva 95/46/CE.

- como assegurar a atualização e a exatidão dos dados que não estão ligados à fonte primária;
- como assegurar que a utilização dos dados pessoais não ultrapassa as funcionalidades previstas para a finalidade inicial da publicação;
- como assegurar a eliminação oportuna dos dados se a publicação de dados pessoais estiver prevista apenas para um período de tempo determinado[34];
- como exercer os direitos dos indivíduos relativamente aos dados pessoais disponibilizados para reutilização (incluindo o direito a exigir a retificação, atualização ou supressão).

7.9. Restrições jurídicas e/ou técnicas à reutilização

Por vezes, a legislação ou a conceção técnica dos sistemas impõe restrições a operações de tratamento específicas ou estabelece outras garantias que restringem a utilização dos registos públicos (nomeadamente, limitando a possibilidade de descarregar todo o conteúdo do registo ou limitando as capacidades de pesquisa, por exemplo, com base no nome e apelido de um indivíduo). Neste caso, a reutilização só deve ser em princípio autorizada se estiver de acordo com estas restrições e condições específicas.

Neste contexto, é importante analisar cuidadosamente que medidas – jurídicas e técnicas – poderiam ser aplicaadas para ajudar a dar resposta às preocupações que se colocam no domínio da proteção de dados, incluindo as preocupações descritas na secção 7.8. É particularmente importante considerar o modo como os reutilizadores terão acesso aos dados – por exemplo, através de uma funcionalidade de descarregamento de grandes quantidades de dados ou de uma interface personalizada com capacidades de acesso limitadas e sujeitas a determinadas condições. Neste aspeto, é crucial considerar os controlos de segurança adicionais que serão aplicados, como, por exemplo, um sistema de verificação «Captcha[35]» para evitar o acesso automatizado e minimizar o risco de «colheita» de uma base de dados completa. O recurso a medidas técnicas específicas poderia ajudar a reduzir os casos de utilização abusiva de dados pessoais e de impactos negativos sobre as pessoas em causa, que seriam possíveis se os reutilizadores tivessem um acesso ilimitado e incondicional a bases de dados completas.

Importa referir que, em muitos casos, será necessário assegurar que os reutilizadores só poderão efetuar consultas específicas através de tecnologias destinadas a evitar o descarregamento de grandes quantidades de dados, nomeadamente através de interfaces de programação de aplicações («API») personalizadas. Estas medidas ajudarão a assegurar a proporcionalidade da utilização e a reduzir os riscos de utilização abusiva de bases de dados completas. Além disso, estas interfaces personalizadas também poderão ajudar a assegurar que os dados estarão sempre atualizados e que deixarão de estar disponíveis através da API quando o organismo do setor público em causa tomar uma decisão nesse sentido. Por outro lado, poderá limitar as formas de reutilização de dados por parte de um reutilizador.

---

[34] Ver, por exemplo, o acórdão do Tribunal de Justiça da União Europeia, de 9 de novembro de 2010, nos processos apensos C-92/09 e C-93/09, *Volker und Markus Schecke GbR/Land Hessen*, n.º 31: «É impossível retirar os dados da Internet após o termo do prazo de dois anos previsto no artigo 3.º, n.º 3, do Regulamento n.º 259/2008».

[35] Um CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*) é um teste de desafio/resposta concebido para distinguir entre humanos e programas automatizados. Um CAPTCHA distingue um ser humano de um computador propondo uma tarefa que é de fácil execução para os seres humanos, mas é mais difícil para os atuais programas informáticos.

7.10. Exatidão, atualizações e eliminação

Outra questão específica prende-se com o que acontece se os dados pessoais forem publicados ou disponibilizados de outra forma ao público apenas por tempo limitado. O artigo 6.º, n.º 1, alínea e), da Diretiva 95/46/CE dispõe que os dados pessoais serão conservados de forma a permitir a identificação das pessoas em causa apenas durante o período necessário para a prossecução das finalidades para que foram recolhidos ou para que são tratados posteriormente. O considerando 18 da Diretiva ISP também estabelece que, caso os «serviços responsáveis decidam deixar de colocar à disposição determinados documentos, devem tornar pública tal decisão, em tempo oportuno e por meios eletrónicos sempre que possível».

Porém, é difícil e, por vezes, impossível assegurar a eliminação ou a retirada dos dados depois de terem sido publicados e disponibilizados para reutilização.

Nesta matéria, uma solução – que apenas resolveria parcialmente o problema – passaria por não disponibilizar os dados em formato descarregável, mas somente através de uma API personalizada e sujeitos a certas restrições e medidas de segurança, tal como referido anteriormente.

VIII. Dados da investigação

É importante estabelecer aqui uma distinção entre a publicação de dados anonimizados, por um lado (ver secção VI), e acesso limitado, por outro. É evidente que a agenda dos dados abertos depende da disponibilidade pública dos dados. No entanto, grande parte da investigação (em especial, a investigação científica, tanto para fins comerciais como não comerciais, mas também outro tipo de investigação) tem lugar através da divulgação dos dados no seio de uma comunidade fechada, ou seja, em que um número finito de investigadores ou instituições tem acesso aos dados, sendo possível restringir a divulgação ou utilização posterior dos dados e garantir a sua segurança.

O acesso limitado é particularmente importante para o tratamento de dados pessoais (frequentemente apresentados sob forma de pseudónimos[36]) provenientes de material sensível e quando exista um risco significativo de reidentificação. Embora a divulgação com acesso limitado não esteja isenta de riscos, estes são menores e é mais fácil minimizá-los quando os dados são divulgados no seio de uma comunidade fechada que funciona com base em regras previamente estabelecidas.

Um problema que se coloca frequentemente aos utilizadores de dados para fins de investigação é o facto de, por um lado, quererem dados ricos, granulares e suficientemente utilizáveis para os fins pretendidos e, por outro, desejarem evitar a reidentificação dos indivíduos. Num extremo do espetro, os dados sob pseudónimos individuais (por exemplo, simplesmente codificados com chave) podem ser muito úteis para os investigadores devido à sua granularidade ao nível individual e ao facto de ser relativamente fácil combinar registos sob pseudónimo de diferentes fontes. No entanto, existe também um elevado risco de reidentificação: a possibilidade de associar vários conjuntos de dados (sob pseudónimo ou não) ao mesmo indivíduo por ser um precursor da identificação ou permitir a identificação direta.

Por conseguinte, é necessário um exame mais atento e uma cautela adicional antes de publicar ou disponibilizar para reutilização conjuntos de dados sob pseudónimo. Em regra, quanto mais

[36] Ver, mais uma vez, o Parecer 4/2007 sobre o conceito de dados pessoais, adotado em 20.6.2007 (WP 136), especialmente nas páginas 12-21 (em que se discute os conceitos de «dados sob pseudónimo», «dados codificados com chave» e «dados anónimos» nas páginas 18-21). A questão da informação «relativa a» uma pessoa singular é discutida nas páginas 9-12. Também é relevante o facto, tal como referido na página 3, de o GT29 estar atualmente a trabalhar na elaboração de orientações adicionais sobre técnicas de anonimização.

pormenorizados, associáveis e individualizados forem os dados, mais limitado e controlado deve ser o acesso aos mesmos. Quanto mais agregados e menos associáveis forem os dados, maior é a probabilidade de poderem ser publicados e disponibilizados para reutilização sem riscos significativos.

Trata-se de uma área complexa e em evolução, pelo que não seria adequado excluir categoricamente a publicação e reutilização de todos os conjuntos de dados que não atingissem o elevado grau de «anonimização» descrito na secção VI. Não obstante, e embora se imponha uma análise casuística e uma avaliação cuidadosa, o GT29 considera que, em regra, a publicação, ao abrigo da Diretiva ISP, de conjuntos de dados ao nível individual ou de outros conjuntos de dados que coloquem um risco significativo de reidentificação nem sempre será adequada.

Além disso, importa salientar que, se alguns desses conjuntos de dados forem, ainda assim, publicados e disponibilizados após uma avaliação cuidadosa dos riscos e benefícios, a divulgação e qualquer reutilização posterior terão de estar em plena conformidade com a legislação sobre proteção de dados (ver secção VII). Tal resulta do facto de estes dados, não obstante as medidas (por vezes, muito significativas) adotadas para reduzir os riscos de reidentificação, continuarem a ser considerados dados pessoais.

IX. Arquivos históricos

Os arquivos históricos e os museus também têm características específicas que exigem garantias específicas. Em muitos casos, e dependendo de fatores como a antiguidade e a sensibilidade dos dados e o contexto em que foram recolhidos, existem outras opções – tais como autorizar apenas um acesso restrito e sujeito a obrigações de confidencialidade – que poderão ser mais adequadas do que a digitalização e disponibilização dos dados para reutilização através da Internet, sem restrições.

Relativamente aos arquivos, é igualmente importante salientar que, embora a sensibilidade dos dados diminua geralmente com a passagem do tempo, a publicação inoportuna de registos com várias décadas poderia ainda ter um efeito muito prejudicial não só sobre a pessoa diretamente afetada, como também sobre outras pessoas, como os membros da sua família, ou os seus descendentes. O risco é particularmente elevado quando estão em causa dados extremamente sensíveis. Por exemplo, a publicação de registos criminais continuaria a estigmatizar o indivíduo e a dificultar a sua reabilitação. A informação de que uma pessoa já falecida era um agente secreto ou colaborador de um regime opressor, um pedófilo, um criminoso, sofria de uma doença mental geradora de estigmatização ou sofria de uma doença hereditária também poderá ter um impacto negativo sobre a família (por exemplo, o cônjuge sobrevivo, os filhos ou outros descendentes) do falecido. As amostras de ADN de pessoas falecidas, que são por vezes conservadas nos arquivos dos hospitais públicos, também poderão exigir proteção por razões semelhantes. Consequentemente, estas informações, ainda que sejam relativas a pessoas já falecidas, poderão exigir proteção ao abrigo da legislação sobre proteção de dados e/ou de outra legislação de defesa dos direitos fundamentais, conforme os casos.

Muitos Estados-Membros aprovaram leis específicas que regulam o acesso aos arquivos nacionais, aos arquivos de períodos históricos recentes de especial interesse (como os arquivos de provas de colaboração com regimes opressores) e a processos conservados pelas autoridades judiciais.[37] Estas

[37] Outros exemplos incluem os arquivos dos registos civis, que contêm, em alguns Estados-Membros, informações como causa do óbito, mudança de sexo e nome do parceiro (de onde é possível inferir a orientação sexual), ou o facto de uma pessoa ter sido adotada. O acesso a estes arquivos também está sujeito a condições específicas.

leis exigem frequentemente a adoção de medidas de segurança adequadas e a imposição de restrições ao acesso, bem como outras garantias que visam o equilíbrio entre os interesses em jogo e a garantia da acessibilidade a certos dados pessoais para fins de investigação histórica, transparência e investigação jornalística, assegurando simultaneamente que a divulgação, quando necessária, seja limitada a fim de não prejudicar a vida privada e familiar e a dignidade dos interessados.

Quanto à «limitação da finalidade», importa referir que os arquivos históricos armazenam geralmente informações para fins de investigação histórica. Estes fins são diferentes dos fins para que os dados foram originalmente recolhidos. Os materiais que acabarão por integrar os acervos arquivísticos foram inicialmente criados para fins administrativos específicos por diferentes organismos do setor público. Geralmente, decorrido um certo período de tempo, quando os documentos já não são necessários para os fins administrativos iniciais, é realizado um processo de seleção e os documentos que forem considerados de valor «histórico» são transferidos para os arquivos históricos. Aqui, a questão que se coloca é a de saber para que fins os dados pessoais armazenados nos arquivos devem ser disponibilizados para reutilização. Neste contexto, é importante realizar uma avaliação cuidadosa, na qual sejam ponderados não só os potenciais benefícios da disponibilização do material arquivístico para reutilização, mas também o potencial impacto sobre os direitos, as liberdades e a dignidade das pessoas em questão.

De um modo geral, pode concluir-se que, embora a digitalização de certos registos que contêm dados pessoais e a sua disponibilização para reutilização possam ser adequadas em algumas situações e alguns dados possam ser publicados sob forma anonimizada, noutros casos é fundamental impor limites à divulgação e reutilização de dados pessoais e estabelecer medidas de segurança adequadas para proteger esses dados. Deve ser realizada uma avaliação minuciosa do impacto na proteção de dados para garantir que os acervos arquivísticos apenas serão disponibilizados para reutilização se forem excluídos quaisquer impactos negativos sobre as pessoas em questão ou se esses riscos forem reduzidos a um mínimo aceitável. O setor dos arquivos também poderia ponderar a elaboração de códigos de conduta ou alterar os códigos existentes para explicar boas práticas.

X. Licenciamento de dados pessoais para reutilização

10.1. Disposições relevantes da Diretiva ISP

O considerando 15 da Diretiva ISP estabelece que «garantir a clareza e a disponibilização ao público das condições de reutilização dos documentos do setor público constitui um requisito prévio ao desenvolvimento de um mercado da informação à escala comunitária. Assim, todas as condições aplicáveis à reutilização dos documentos do setor público devem ser claramente apresentadas aos potenciais reutilizadores. Os Estados-Membros deverão incentivar a criação de índices dos documentos disponíveis, acessíveis em linha, se for caso disso, para promover e facilitar os pedidos de reutilização».

O considerando 26 da Alteração ISP estabelece ainda que «em relação à reutilização feita dos documentos, os organismos do setor público podem impor condições, se adequado através de uma licença...» e que «os Estados-Membros deverão, se adequado, encorajar a utilização de formatos abertos legíveis por máquina».

Além disso, o artigo 8.º, n.º 1, dispõe que os «organismos do setor público podem autorizar a reutilização sem condições ou podem impor condições, se adequado através de uma licença. Essas condições não devem restringir desnecessariamente as possibilidades de reutilização e não devem ser utilizadas para limitar a concorrência.»

10.2. Licenciamento e proteção de dados

As licenças são um elemento central do regime de ISP. Podem também afetar o modo como os dados pessoais são tratados e devem ser uma das garantias impostas quando são disponibilizados dados pessoais (ou dados anonimizados obtidos a partir de dados pessoais) para reutilização. As licenças não dispensam o cumprimento da legislação sobre proteção de dados, mas a inclusão de uma cláusula sobre esta matéria nas condições da licença ajudaria a assegurar esse cumprimento, reforçando a «obrigatoriedade». Uma cláusula deste tipo seria igualmente importante em termos de sensibilização, na medida em que recordaria aos reutilizadores as suas obrigações na qualidade de responsáveis pelo tratamento de dados.

No que respeita ao teor das licenças, é útil distinguir entre dois cenários diferentes.

10.3. Condições da licença para conjuntos de dados anonimizados

Em primeiro lugar, estando em causa dados anonimizados (ou seja, conjuntos de dados que já não contêm dados pessoais), as condições das licenças devem:

- reiterar que os conjuntos de dados têm de ser anonimizados;
- proibir os titulares das licenças de reidentificar quaisquer pessoas[38];
- proibir os titulares das licenças de utilizar os dados para tomar qualquer medida ou decisão em relação às pessoas em questão; e
- impor ao titular da licença a obrigação de notificar a entidade licenciadora caso seja detetado que as pessoas podem ser ou foram reidentificadas.

Ao invés de serem incluídas na licença, estas condições também poderiam assumir a forma de um aviso aos reutilizadores que figurasse, de forma proeminente, no portal de dados abertos. No entanto, é preferível promover a primeira opção porque apresenta a vantagem de ter eficácia contratual.

Retirada de conjuntos de dados comprometidos

Todos os cibernautas, incluindo as próprias pessoas em causa, devem ter a possibilidade de alertar a entidade licenciadora para o facto de ter ocorrido ou poder ocorrer uma reidentificação. Nos casos em que a entidade licenciadora detete um risco acrescido de identificação, a licença deve prever um procedimento que permita a esta entidade «retirar» o conjunto de dados «comprometido». Por outras palavras, a cláusula sobre proteção de dados deve conferir à entidade licenciadora o direito de suspender ou cancelar o acesso aos dados (por exemplo, o direito de desativar a API ou de retirar o ficheiro da plataforma). A entidade licenciadora deve envidar todos os esforços razoáveis para exigir que todos os reutilizadores eliminem a totalidade ou partes dos conjuntos de dados que foram comprometidos (que se tornaram reidentificáveis), nomeadamente colocando avisos num local proeminente em sítios *Web* como os portais de dados abertos e fóruns/listas de endereços eletrónicos/redes sociais acedidos por grupos ou indivíduos que provavelmente estarão a reutilizar os dados. A exigência de registo poderá ser o meio mais eficaz de retirar conjuntos de dados, mas esta

[38] Poderiam ser previstas algumas exceções, nomeadamente em casos de testes de reidentificação realizados de boa-fé. Porém, mesmo nestes casos, os resultados dos testes devem ser transmitidos ao responsável pelo tratamento e ao organismo do setor público em causa e os dados reidentificados não devem ser publicados nem ser objeto de uma divulgação mais vasta por outros meios.

solução não deve ser encorajada se implicar a recolha de novos dados pessoais dos reutilizadores e se tiver o efeito geral de desincentivar a utilização de sítios *Web* de ISP e outros serviços.

10.4. Condições das licenças para dados pessoais

Quando é concedida uma licença para dados pessoais, é necessário definir os limites da utilização desses dados. Neste caso, a principal preocupação consiste em assegurar que qualquer reutilização se restringirá ao que for compatível com as finalidades para que os dados foram inicialmente recolhidos[39]. Para tal, é necessário, pelo menos, especificar claramente nas condições da licença as finalidades para que os dados foram inicialmente publicados e dar uma indicação das situações que seriam e que não seriam consideradas como utilização compatível de dados pessoais.

Importa salientar, porém, que estas condições não devem «restringir desnecessariamente as possibilidades de reutilização» (artigo 8.º, n.º 1, da Alteração ISP). Por este motivo, as condições genéricas das licenças-tipo abertas nem sempre serão adequadas, podendo ser necessário criar licenças específicas para certos tipos de dados pessoais ou utilizar modelos que poderiam ser adaptados.

Atualmente, algumas licenças-tipo abertas (como a licença governamental aberta do Reino Unido) excluem os dados pessoais – que, de acordo com as condições estabelecidas, não são de forma alguma licenciados.

10.5. Imposição de sanções em caso de reidentificação ou utilização incompatível

Depois de os dados terem sido publicados ao abrigo de uma licença – como uma licença governamental aberta – poderá ser difícil protegê-los de utilizações incompatíveis posteriores, divulgação ou manter a sua segurança. O controlo da reutilização e a resposta a violações, quer assumam a forma de reidentificação das pessoas em causa ou de utilizações posteriores para uma finalidade incompatível pela entidade licenciadora, são muito importantes neste contexto.

Embora o GT29 reitere o importante papel que os organismos do setor público devem desempenhar, salienta igualmente que, se um reutilizador recolher dados pessoais através de um processo de reidentificação, é muito provável que se considere que está a proceder a um tratamento ilícito de dados pessoais punível pelas autoridades de proteção de dados. As sanções previstas no Regulamento «Proteção de Dados» proposto incluem multas de elevado valor.

[39] Ver, mais uma vez, o Parecer 3/2013 do GT29 sobre a limitação da finalidade.

XI. Conclusões

Em conclusão, o GT29 reitera que a reutilização de ISP poderá trazer benefícios que conduzam a uma maior transparência e a uma reutilização inovadora das informações do setor público. Porém, a maior acessibilidade das informações daí resultante comporta riscos. A fim de assegurar a proteção da privacidade e dos dados pessoais dos indivíduos, é necessário seguir uma abordagem equilibrada e a legislação sobre proteção de dados tem de ajudar a orientar o processo de seleção dos dados pessoais que podem e não podem ser disponibilizados para reutilização e das medidas a adotar para os proteger.

Independentemente do «princípio da reutilização» formulado na Alteração ISP, a reutilização para fins comerciais ou não comerciais ao abrigo da Diretiva ISP nem sempre será adequada em casos em que as ISP a reutilizar contenham dados pessoais. Não são geralmente dados pessoais, mas sim dados estatísticos obtidos a partir de dados pessoais, que são e que devem ser disponibilizados para reutilização.

Não obstante, em algumas situações, os dados pessoais poderão ser considerados disponíveis para reutilização ao abrigo da Diretiva ISP, sendo adotadas, se necessário, medidas jurídicas, técnicas ou organizativas adicionais para proteger as pessoas em causa. Nestes casos, o GT29 reitera a importância de estabelecer uma base legal sólida para disponibilizar ao público dados pessoais, tomando em consideração as regras sobre proteção de dados aplicáveis, incluindo o princípio da proporcionalidade, da minimização dos dados e da limitação da finalidade. Neste contexto, também é importante salientar, mais uma vez, que quaisquer informações relativas a uma pessoa singular identificada ou identificável, estejam ou não publicamente disponíveis, constituem dados pessoais. Por conseguinte, o acesso e a reutilização de dados pessoais que tenham sido disponibilizados ao público continuam sujeitos à legislação sobre proteção de dados aplicável.

À luz destas considerações, o GT29 recomenda o seguinte:

- o facto de algumas ISP poderem conter dados pessoais deve ser tomado em consideração o mais cedo possível quando estiver a ser ponderada a disponibilização de ISP ao público, em conformidade com os princípios da «proteção de dados desde a conceção» e da «proteção de dados por defeito»;
- neste sentido, o organismo do setor público em causa (ou o legislador, conforme os casos) deve realizar uma avaliação do impacto na proteção de dados antes de disponibilizar para reutilização ISP que contenham dados pessoais (ou antes de adotar uma lei que permita a publicação de dados pessoais e, como tal, os disponibilize potencialmente para reutilização); deve ser realizada uma avaliação do impacto na proteção de dados nas situações em que esteja em causa a disponibilização para reutilização de conjuntos de dados anonimizados obtidos a partir de dados pessoais;
- quando forem anonimizados conjuntos de dados, é essencial avaliar o risco de reidentificação e dispor de uma boa prática para realizar testes de reidentificação;
- o resultado da avaliação poderá ajudar a identificar garantias adequadas para minimizar os riscos, incluindo, por exemplo, medidas técnicas, jurídicas e administrativas como a inclusão de condições adequadas na licença e medidas técnicas para evitar o descarregamento de grandes quantidades de dados, e técnicas de anonimização adequadas; poderá também levar à decisão de não publicar e/ou disponibilizar para reutilização;
- a licença de reutilização de ISP deve incluir uma cláusula de proteção de dados sempre que sejam tratados dados pessoais, nomeadamente nos casos em que sejam disponibilizados para reutilização conjuntos de dados anonimizados obtidos a partir de dados pessoais;

- sempre que a avaliação do impacto na proteção de dados conclua que uma licença aberta não é suficiente para responder aos riscos em matéria de proteção de dados, os organismos do setor público não devem disponibilizar dados pessoais ao abrigo da Diretiva ISP. (No entanto, o organismo do setor público poderá, ainda assim, no âmbito da sua discricionariedade, ponderar a reutilização fora das condições e do âmbito da Diretiva ISP e poderá também exigir aos requerentes que demonstrem ter dado uma resposta adequada aos riscos para a proteção de dados pessoais e que irão tratar os dados em conformidade com a legislação sobre proteção de dados aplicável);
- os organismos do setor público devem, se for o caso, certificar-se de que os dados pessoais foram anonimizados e de que as condições da licença proíbem expressamente a reidentificação de indivíduos e a reutilização de dados pessoais para fins suscetíveis de afetar as pessoas em causa;
- por último, os Estados-Membros devem ainda ponderar a criação e o apoio de redes de conhecimento/centros de excelência e viabilizar, deste modo, a partilha de boas práticas relacionadas com anonimização e dados abertos.

Feito em Bruxelas, em 5 de junho de 2013

Pelo Grupo de Trabalho
O Presidente
Jacob KOHNSTAMM