Os Web Crawlers, Scrapping e serviços de raspagem web ficam ilegais com a GDPR e Lei Geral de Proteção de Dados?

Não é novidade que a LGPD, inspirada pela GDPR, entra em vigor em agosto de 2020. As normas mencionadas são bem claras quando o assunto é responsabilidade dos atores que recebem os dados no âmbito da expectativa do titular, com seu consentimento ou amparados por uma das premissas que permitem o tratamento.

A legislação estabelece, por exemplo, que o controlador poderá prever ou ceder dados do titular ao processador, orientando este para que siga as orientações e principalmente ofereça níveis aceitáveis de proteção de dados.

Por outro lado, nem sempre tudo é tão simétrico e os agentes de tratamento são claramente definidos. Em muitos casos, os dados chegam a terceiros de forma não autorizada pelo controlador, ou até mesmo sem que este saiba. Ferramentas que instrumentalizam técnicas de web scrapping e crawling ou mesmo automatizam pesquisas são constantemente utilizadas para tratamento (coleta) de dados de modo nem sempre convencional.

Nestas técnicas, hoje, estão assentados o núcleo de muitos negócios envolvendo empresas de marketing digital, inteligência de dados, background screening e scoring, sendo hoje também a base para inúmeros negócios digitais e startups.

Nesta modalidade, via de regra, o controlador não disponibilizou ao agente de tratamento um webservice, uma API, uma exportação ou qualquer meio “legal”, “previsto”, “dimensionado” ou “regular” para que terceiros acessem dados

de seus titulares.

Uma pessoa física ou jurídica acessa alguma interface de exibição de dados do cliente e coleta os mesmos, iniciando um novo ciclo de tratamento. E que nem se argumente que estes dados “foram tornados públicos”, logo, fora do escopo das normas de proteção de dados, pois quem torna um dado público é o titular e muitos crawlers operam inclusive em ambiente logado ou contornando barreiras tecnológicas. Assim, cada caso deverá ser analisado em minúcia.

Mas as atividades seriam ilegais? Neste ponto vale destacar que o artigo 17 da GDPR é claro ao prever que o titular dos dados tem inúmeros direitos, inclusive o direito a exclusão, quando os dados forem tratados ilicitamente. A própria consideranda 39 da norma, deixa claro que os dados pessoais deverão ser tratados de modo a não serem utilizados por pessoas não autorizadas. Não bastasse, a definição de “Violação de dados pessoais” também engloba um acesso não autorizado ou ilícito.

O mesmo se repete no art. 46 da LGPD, que estabelece que os agentes de tratamento devem adotar medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito.

Não bastasse, na LGPD, está previsto no artigo 18 o direito do titular dos dados à eliminação de dados tratados em desconformidade com a Legislação, estabelecendo a lei nacional no seu artigo 44 o conceito de tratamento irregular de dados, incluindo quando deixar de observar a legislação ou quando não fornecer a segurança que o titular dele pode esperar, consideradas as circunstâncias relevantes, entre as quais o modo que é realizado.

Neste sentido, a menos que exista uma relação clara com o

controlador e uma premissa válida de tratamento, o crawling pode ser considerado irregular. Na maioria dos casos o consentimento será a hipótese válida para estas operações, o que se sabe não é tarefa fácil de se conseguir. A nós, parece muito temerário e forçoso tentar demonstrar que um serviço baseado em crawling opera com base em interesse legítimo, considerando que muitas vezes o titular e até mesmo o controlador não sabem da atividade (Este, que poderá detectar mediante monitoramento de requisições a seu serviço).

Na Europa, serviços de crawling já estabelecem em seus termos em que há raspagem de dados apenas com o consentimento. Para estas empresas, recomenda-se analisar quais dados tratam e compõe sua base de dados e que foram objeto de raspagem. Após, é importante avaliar qual a premissa válida para tratar estes dados que foram empregados. Não estando abrangido por uma premissa válida e sem consentimento, a recomendação é exclusão ou anonimização dos dados. E diante de qualquer nova iniciativa que envolva raspagem, a recomendação básica para minimizar riscos é um contato com o controlador para compreender sua expectativa e a do próprio titular dos dados pessoais.

Com efeito, conquanto não possam ser considerados ilegais em todos os casos, que deverão ser cuidadosamente analisados em suas peculiaridades, é evidente que existem cautelas necessárias para empresas que fazem este tipo de tratamento de dados, considerando que controladores que tratem dados legalmente, de modo até a assegurarem os princípios e diretrizes das normas e minimizar riscos de problemas com titulares, poderão não só dificultar o acesso não convencional, mas periciar seus sistemas, identificar coletas indevidas e processar os referidos serviços.

E por falar nas medidas de controladores e operadores, as recomendações para estes, caso passem por uma ação legal ou intervenção administrativa diante de um tratamento não autorizado de dados que lhes foram confiados é, justamente,

revisar não só como dados são armazenados na base, mas como são expostos em suas interfaces de acessibilidade, revisando os processos para que possam demonstrar claramente que adotaram medidas técnicas e organizativas possíveis para redução da formação de ciclos de vida indevidos e desconhecidos, estando preparados para apresentar estas medidas a qualquer momento.

Do mesmo modo, é muito importante, diante de uma violação de dados a partir de tratamentos feitos por terceiros sem relação com o controlador ou operador que estes tenham procedimentos claros para demonstrar que não realizaram o referido tratamento, considerando o disposto no art. 82 da GDPR e art. 43, I da LGPD, que é claro ao prever que os referidos agentes de tratamento são isentos de responsabilidade se provarem que não são de modo algum responsáveis ou não realizaram o tratamento indevido. Não se pode desconsiderar, em alguns casos (e cada caso, como dito, terá suas peculiaridades), controladores e operadores sejam "tão vítimas", quanto os próprios titulares dos dados.

Não é novidade que a LGPD, inspirada pela GDPR, entra em vigor em agosto de 2020. As normas mencionadas são bem claras quando o assunto é responsabilidade dos atores que recebem os dados no âmbito da expectativa do titular, com seu consentimento ou amparados por uma das premissas que permitem o tratamento.

A legislação estabelece, por exemplo, que o controlador poderá prever ou ceder dados do titular ao processador, orientando este para que siga as orientações e principalmente ofereça níveis aceitáveis de proteção de dados.

Por outro lado, nem sempre tudo é tão simétrico e os agentes de tratamento são claramente definidos. Em muitos casos, os dados chegam a terceiros de forma autorizada pelo

controlador, ou até mesmo sem que este saiba. Ferramentas que instrumentalizam técnicas de web scrapping e crawling ou mesmo automatizam pesquisas são constantemente utilizadas para tratamento (coleta) de dados de modo nem sempre convencional.

Nestas técnicas, hoje, estão assentados o núcleo de muitos negócios envolvendo empresas de marketing digital, inteligência de dados, background screening e scoring, sendo hoje também a base para inúmeros negócios digitais e startups.

Nesta modalidade, via de regra, o controlador não disponibilizou ao agente de tratamento um webservice, uma API, uma exportação ou qualquer meio “legal”, “previsto”, “dimensionado” ou “regular” para que terceiros acessem dados de seus titulares.

Uma pessoa física ou jurídica acessa alguma interface de exibição de dados do cliente e coleta os mesmos, iniciando um novo ciclo de tratamento. E que nem se argumente que estes dados “foram tornados públicos”, logo, fora do escopo das normas de proteção de dados, pois quem torna um dado público é o titular e muitos crawlers operam inclusive em ambiente logado ou contornando barreiras tecnológicas. Assim, cada caso deverá ser analisado em minúcia.

Mas as atividades seriam ilegais?
Neste ponto vale destacar que o artigo 17 da GDPR é claro que o titular dos

dados tem inúmeros direitos, inclusive o direito a exclusão, quando os dados forem tratados ilicitamente. A própria consideranda 39 da norma, deixa claro que os dados pessoais deverão ser tratados de modo a não serem utilizados por pessoas não autorizadas. Não bastasse, a definição de “Violação de dados pessoais” também engloba um acesso não autorizado ou ilícito.

O mesmo se repete no art. 46 da LGPD, que estabelece que os agentes de tratamento devem adotar medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito.

Não bastasse, na LGPD, está previsto no artigo 18 o direito do titular dos dados à eliminação de dados tratados em desconformidade com a Legislação, estabelecendo a lei nacional no seu artigo 44 o conceito de tratamento irregular de dados, incluindo quando deixar de observar a legislação ou quando não fornecer a segurança que o titular dele pode esperar, consideradas as circunstâncias relevantes, entre as quais o modo que é realizado.

Neste sentido, a menos que exista uma relação clara com o controlador e uma premissa válida de tratamento, o

crawling pode ser considerado irregular. Na maioria dos casos o consentimento
será a hipótese válida para estas operações, o que se sabe não é tarefa fácil
de se conseguir. A nós, parece muito temerário e forçoso tentar demonstrar que
um serviço baseado em crawling opera com base em interesse legítimo,
considerando que muitas vezes o titular e até mesmo o controlador sabem da
atividade (Este, que poderá detectar mediante monitoramento de requisições a
seu serviço).

Na Europa, serviços de crawling já estabelecem em seus termos que há raspagem de dados apenas com o consentimento. Para estas empresas, recomenda-se analisar quais dados tratam e compõe sua base de dados e que foram objeto de raspagem. Após, é importante avaliar qual a premissa válida para tratar estes dados que foram empregados. Não estando abrangido por uma premissa válida e sem consentimento, a recomendação é exclusão ou anonimização dos dados. E diante de qualquer nova iniciativa que envolva raspagem, a recomendação básica para minimizar riscos é um contato com o controlador para compreender sua expectativa e a do próprio titular dos dados pessoais.

Com efeito, conquanto não possam
ser considerados ilegais em todos os casos, que deverão ser cuidadosamente
analisados em suas peculiaridades, é evidente que existem cautelas necessárias
para empresas que fazem este tipo de tratamento de dados, considerando que
controladores que tratem dados legalmente, de modo até a assegurarem os
princípios e diretrizes das normas e minimizar riscos de

problemas com
titulares, poderão não só dificultar o acesso não convencional, mas periciar
seus sistemas, identificar coletas indevidas e processar os referidos serviços.

E por falar nas medidas de controladores e operadores, as recomendações para estes, caso passem por uma ação legal ou intervenção administrativa diante de um tratamento não autorizado de dados que lhes foram confiados é, justamente, revisar não só como dados são armazenados na base, mas como são expostos em suas interfaces de acessibilidade, revisando os processos para que possam demonstrar claramente que adotaram medidas técnicas e organizativas possíveis para redução da formação de ciclos de vida indevidos e desconhecidos, estando preparados para apresentar estas medidas a qualquer momento.

Do mesmo modo, é muito importante, diante de uma violação de dados a partir de tratamentos feitos por terceiros sem relação com o controlador ou operador que estes tenham procedimentos claros para demonstrar que não realizaram o referido tratamento, considerando o disposto no art. 82 da GDPR e art. 43, I da LGPD, que é claro a prever que os referidos agentes de tratamento são isentos de responsabilidade se provarem que não são de modo algum responsáveis ou não realizaram o tratamento indevido. Não se pode desconsiderar, em alguns casos (e cada caso, como dito, terá suas peculiaridades), controladores e operadores sejam “tão vítimas”, quanto os próprios titulares dos dados.