

57

DELIBERAÇÃO N.º 929/2014

I. Pedido

O Decreto-Lei n.º 226/2009, de 14 de setembro, no uso da autorização legislativa concedida pela Lei n.º 35/2009, de 14 de julho, estabeleceu que o Recenseamento de 2011 deveria ser a última operação censitária a realizar em Portugal com recurso a um inquérito exaustivo junto de toda a população.

Deste modo, os Censos 2021 deverão obedecer a um novo modelo censitário, centrado na utilização de fontes administrativas, substituindo, assim, o modelo tradicional baseado em inquéritos junto da população, de forma a permitir a redução dos custos e da carga estatística, assim como uma maior frequência na informação censitária a disponibilizar.

Para aferir a viabilidade deste novo modelo, que tem como pressuposto o acesso aos dados pessoais detidos por vários organismos da administração pública, o Instituto Nacional de Estatística (INE) necessita de realizar testes previamente à publicação da legislação que determinará a realização da operação censitária no ano de 2021.

Com este fim o INE dirigiu-se à Comissão Nacional de Protecção de Dados (CNPd) no sentido de ser encontrada uma solução que, concretizando o objetivo do novo modelo censitário, não afetasse o direito à protecção de dados dos cidadãos. Neste contexto, foi criado um mecanismo de estreita colaboração entre as duas instituições de modo a encontrar as soluções técnicas que permitissem atingir estes desideratos.

Veio agora o INE submeter à apreciação da CNPD as soluções técnicas encontradas para a realização do estudo de viabilidade dos Censos 2021 por recurso a fontes administrativas.

27



A integração da informação proveniente de diferentes fontes de dados implica a ligação dos vários ficheiros de diversas entidades através de identificadores numéricos e, se necessário, de outros dados.

O pressuposto central é o de que os dados individuais não sejam suscetíveis de identificação. Deste modo, as diferentes fontes administrativas, responsáveis pelos tratamentos de dados pessoais, aplicarão, na fonte, um sistema de encriptação de dados por aplicação de um *hash* com o algoritmo SHA256, recebendo o INE os ficheiros com a informação já anonimizada.

Para a integração informática dos vários ficheiros, serão utilizados identificadores numéricos – número de identificação (NIC), fiscal (NIF) e da segurança social (NISS), aos quais será aplicado o supracitado *hash* antes envio ao INE.

Como estes identificadores não constam em todas as fontes administrativas, para permitir a ligação dos ficheiros cada entidade deverá incluir outros elementos que permitam garantir que a informação respeita à mesma unidade.

Assim, “localidade” e o “código postal” servirão apenas para apoiar a decisão de emparelhamento dos registos sobre os quais subsistam dúvidas. O mesmo papel deverá ter o “nome”, o qual será abreviado antes do envio ao INE de forma a não permitir a identificação dos indivíduos. A “profissão” e a “atividade económica” são variáveis a transmitir ao INE de acordo com um nível de agregação que impossibilite também a identificação dos cidadãos.

Os dados administrativos necessários, provenientes de nove entidades de cinco ministérios, são os seguintes:

1. Variáveis necessárias para a ligação da informação proveniente das diferentes fontes de dados, a fornecer por todas as entidades responsáveis pela gestão de fontes administrativas:

- 1.1. NIC, encriptado antes do envio ao INE (*)
- 1.2. NISS, encriptado antes do envio ao INE (*)
- 1.3. NIF, encriptado antes do envio ao INE) (*)
- 1.4. Nome: 3 primeiras letras do 1º nome + 3 últimas letras do último nome
- 1.5. Localidade
- 1.6. Código Postal (7 dígitos)

(*) Na medida da disponibilidade de cada n.º de identificação, em cada fonte.

2. Variáveis necessárias por entidade

2.1. Ministério da Justiça

Instituto dos Registos e do Notariado: Base de Dados de Identificação Civil (BDIC): sexo, data de nascimento, estado civil legal, naturalidade (país, município, freguesia), nacionalidade (país).

2.2. Ministério da Administração Interna

Serviço de Estrangeiros e Fronteiras (SEF): sexo, data de nascimento, estado civil legal, naturalidade (país, município, freguesia), nacionalidade (país), profissão (desagregada pela CPP – Classificação Portuguesa de Profissões, versão de 2010).

2.3. Ministério da Educação e Ciência

Direção-Geral de Estatísticas da Educação e Ciência (DGEEC): sexo, data de nascimento, nacionalidade (país), frequência de ensino, ano de escolaridade, nível de ensino completo, curso superior, local de estudo (município, freguesia).

2.4. Ministério da Solidariedade, Emprego e Segurança Social

a) Instituto de Informática - Base de Dados da Segurança Social: sexo, data de nascimento, estado civil legal, companheiro em união de facto, naturalidade (país, município, freguesia), nacionalidade (país), local de trabalho (país, município, freguesia), condição perante a atividade económica (empregado, reformado), situação na profissão (conta de outrem, conta própria, etc.), atividade económica (desagregada pela CAE - Classificação de Atividades Económicas, rev.3).

b) Gabinete de Estratégia e Planeamento - Relatório Único / Quadros de Pessoal: sexo, data de nascimento, nível de ensino, local de trabalho (país, município, freguesia), profissão (desagregada pela CPP), n.º horas trabalhadas, atividade económica (desagregada pela CAE), n.º pessoas da empresa.

c) Instituto do Emprego e Formação Profissional: sexo, data de nascimento, nacionalidade (país), nível de ensino, condição perante a atividade económica (desempregados à procura de novo emprego /1º emprego), procura de emprego – tempo, disponibilidade para trabalhar, profissão dos desempregados à procura de novo emprego (desagregada pela CPP).

2.5. Ministério das Finanças

a) Autoridade Tributária – Imposto sobre o Rendimento de Pessoas Singulares (IRS): sexo, data de nascimento, estado civil legal, companheiro em união de facto, naturalidade (país, município, freguesia), nacionalidade (país), fonte de rendimento (trabalho, reforma, rendas, etc.), composição do agregado familiar (NIF's encriptados), número de dependentes.

b) Autoridade Tributária – Imposto Municipal sobre Imóveis (IMI): tipo de edifício, tipo de utilização, n.º de pisos, elevador, n.º de alojamentos, época de construção, necessidade de reparações, tipo de alojamento, água canalizada, retrete, esgotos, área útil, n.º de divisões, estacionamento, entidade proprietária.

c) Autoridade Tributária – Registo de contratos de arrendamento: área útil, n.º divisões, tipo de contrato de arrendamento, época do contrato, escalão de renda.

d) Caixa Geral de Aposentações (CGA): sexo, data de nascimento, local de trabalho (país, município, freguesia), condição perante a atividade económica (empregado, reformado), atividade económica (desagregada pela CAE).

e) Direção-Geral de Proteção Social aos Trabalhadores em Funções Públicas (ADSE): sexo, data de nascimento, estado civil legal, companheiro em união de facto, atividade económica (desagregada pela CAE).

O INE indicou um conjunto de medidas de segurança física e lógica, das quais se destacam, por terem relevância para proteção de dados, as seguintes :

1. Medidas de segurança física

1.1. Os equipamentos que alojam a informação estão localizados no *Datacenter*, protegido por utilização de sistemas de extinção de incêndio, alarmes de intrusão e controlo de acessos por cartões com *PIN*. O *Datacenter* e os equipamentos são monitorizados 24 horas por dia/365 dias por ano.

1.2. Todos os profissionais, funcionários ou prestadores de serviços, estão vinculados à obrigação de confidencialidade da informação a que tenham acesso no exercício das suas atribuições, a qual constitui segredo profissional, mesmo após o termo das respetivas funções.

2. Medidas de segurança lógica

2.1. A segurança lógica é garantida por uma arquitetura seccionada por zonas (rede interna, internet, *dmz* internet, *dmz* interna, etc.), por soluções de controlo de fronteiras (*firewalls*), gestão de identidades, autenticação e gestão de privilégios (LDAP), registo de operações (*logs*) e monitorização das comunicações e serviços aplicacionais.

2.2. O sistema de *backup* utilizado tem ciclos de rotação que permitem o rastreamento histórico e o acesso está sujeito a medidas de segurança reforçadas.



2.3. As comunicações, devidamente encriptadas, serão efetuadas por VPN (*Virtual Private Network*).

2.4 As transmissões eletrónicas de dados das fontes administrativas serão efetuadas em canal de comunicação seguro, com utilização de credenciais de autenticação, em circuitos dedicados.

II. Análise

A forma como se têm desenrolado os CENSOS tem merecido particular atenção por parte da CNPD na medida em que, por definição, a produção de estatística não deveria envolver a identificação de pessoas e que, atento o grau de pormenor e detalhe da informação em causa, os riscos da sua utilização abusiva por terceiros são especialmente gravosos.

O alargado universo de cidadãos abrangidos pelo CENSOS bem como o detalhe da informação a eles atinente, determinam, pois, especiais cuidados na proteção dos dados pessoais e na ulterior anonimização da informação.

Nessa medida, a CNPD acompanhou o desenrolar dos trabalhos relativos à preparação dos CENSOS 2011, tendo emitido, a esse propósito, o Parecer n.º 13/2008, a Deliberação n.º 48/2010 e a Autorização n.º 1308/2011.

O propósito de elaboração dos CENSOS através do recurso às fontes administrativas existentes merece, de resto, toda a cooperação da CNPD, quer pela redução de custos que envolve, quer pela circunstância de não se justificar o recurso a inquéritos exaustivos junto de toda a população quando a informação necessária para a produção estatística existe na Administração Pública com maior fiabilidade e rigor.

Todavia, pela suscetibilidade de o cruzamento da informação existente nos serviços e organismos públicos relativamente a cada um dos cidadãos ser desproporcionado e potencialmente discriminatório, foi necessário encontrar especiais medidas de

segurança que, sem inviabilizar a produção estatística, garantam a privacidade e a protecção de dados pessoais dos cidadãos. Com esse desiderato, a estreita colaboração entre o INE e a CNPD permitiu encontrar uma solução que, atingindo a finalidade pretendida, assegura, do mesmo passo, a protecção de dados pessoais.

Note-se que se qualificam como dados pessoais não só os relativos a pessoa identificada como também a pessoa identificável, entendendo-se, de acordo com o Considerando 26 da Diretiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995, 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.

Contudo, a aplicação nas entidades administrativas detentoras das variáveis necessárias ao CENSOS de um algoritmo de encriptação irreversível – que permite ao INE o cumprimento das suas atribuições e competências no que respeita à realização censitária impossibilitando a identificação dos titulares dos dados – constitui a solução atingida, a qual é garantia da anonimização da informação.

Na medida em que este modelo se mostra viável e posto que não existirá tratamento de dados pessoais na aceção da alínea a) do artigo 3º da LPD, na interpretação dada pelo Considerando 26 da Diretiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995, estão reunidas as necessárias garantias dos cidadãos em matéria de protecção de dados pessoais.

III. Conclusão

1. Face ao exposto, a CNPD delibera no sentido da conformidade do modelo proposto com o regime de protecção de dados pessoais.
2. A CNPD irá acompanhar de perto o estudo de viabilidade relativo ao CENSOS 2021, devendo o INE fornecer-lhe a informação necessária a esse propósito, bem como



reportar-lhe qualquer circunstância que ponha em causa os pressupostos em que assenta a presente deliberação.

Lisboa, 11 de junho de 2014

A handwritten signature in black ink, appearing to read 'Helena Delgado António', is written over a horizontal line.

Helena Delgado António (Relatora)