

Deliberation 2020-027 of February 27, 2020 National Commission for Computing and Liberties Nature of the deliberation:

Opinion Legal status: In force Date of publication on Légifrance: Saturday July 11, 2020 Deliberation No. 2020-027 of February 27, 2020 providing an opinion on a draft automated processing of personal data necessary to carry out a survey called "Chinese immigrants in Paris and the Paris region" (CHIPRE) (request for opinion no. 2214563) The National Commission for Computing and Freedoms, Saisie by the National Institute for Demographic Studies of a request for an opinion concerning a project for the automated processing of personal data necessary to carry out a survey called Chinese immigrants in Paris and the Paris region (Chinese immigrants in the Paris Region or CHIPRE); Having regard to Convention No. 108 of the Council of Europe for the protection of individuals with regard to automatic processing of personal data el; Having regard to Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC; Having regard to law n° 78-17 of January 6, 1978 as amended relating to data processing, files and freedoms, in particular 6° of its article 44; Having regard to decree n° 2019-536 of May 29, 2019 amended taken for the application of law n ° 78-17 of January 6, 1978 relating to data processing, files and freedoms; On the proposal of Mrs. Anne DEBET, commissioner, and after having heard the observations by Mrs. Nacima BELKACEM, Government Commissioner, Issues the following opinion: 1. As a preliminary point, the Commission notes that the survey called Chinese Immigrants in Paris and the Paris Region (CHIPRE) aims to provide data allowing a description of the Chinese population in Ile-de-France and to highlight the crucial aspects of assimilation of immigrants in the French context. 2. In this context, it notes that the planned processing pursues scientific research purposes. In this respect, it considers that the processing falls within the scope of Regulation (EU) 2016/679 of 27 April 2016 referred to above (hereinafter the GDPR) and notes that the National Institute for Demographic Studies (INED) intends to rely on the lawfulness basis mentioned in e) of 1. of Article 6 of the GDPR, namely the performance of a task in the public interest. 3. Insofar as it relates to sensitive data within the meaning of the regulations, the planned processing must be the subject of a prior opinion from the Commission in accordance with the provisions of 6° of article 44 of the law of 6 January 1978 amended, without prejudice to the obligation to formalize, where appropriate, the creation of processing by a specific regulatory act. It recalls that it must be kept informed and informed of any substantial modification affecting the characteristics of the processing and that the impact analysis relating to data protection (DPIA) must be updated. the general context of the survey (CHIPRE) 4. The survey project aims to collect, from a sample of 800 people, data from people born in China, living for

at least three months in Ile-de-France, aged over 18 and under 65 years old at the time of the survey, with the exception of students holding a Chinese government scholarship (in that they are required to return to China immediately upon completion of their studies). 5. The Commission notes that this sample will be made up of ten seeds surveyed, called a convenience sample. The selection criteria for this sample will derive from the conclusions of the field exploration carried out by INED prior to the survey itself, which investigates various networks and relies on its experience during other surveys (associations, traders , family). 6. This survey project is divided into several parts: an exploratory survey, which does not aim to collect personal data but to establish a qualitative pre-knowledge of the structure of the network; the administration of questionnaires; the administration of registers (rosters).7. It also notes that INED plans to conduct some thirty qualitative interviews a posteriori, which are excluded from this referral. The Commission, which notes that these interviews should be audio-recorded, recalls that it will be up to the data controller to ensure that the processing of personal data implemented in this context complies with applicable regulations.8. It also notes that to carry out the planned survey, INED will produce several files, namely: two raw collection files: one corresponding to the responses to the questionnaire, the other containing the register data; the study files: corresponding to the raw files cleared of any reference to coordinates and nominative data. These are then pseudonymised files; the qualitative contact file: this file contains the data of people who have accepted the principle of qualitative interviews; the production and research file: pseudonymized file which will be made available to researchers via the Quetelet network; the file of qualitative interviews.9. Sampling will be carried out using the Network Sampling with Memory (NSM) method: respondents are asked to list a maximum of six members of their entourage or their social network corresponding to the eligibility criteria of the survey and provide information allowing them to be partially identified (the register or roster). 10. The Commission notes that the data collected make it possible to reconstruct the underlying social network of the target population and that the use of the NSM method combines two sampling methods implemented by means of an algorithm programmed by the services of the INED: the research mode which is used to explore the network, selects first and foremost the members of the sociability network of the bridge nodes (respondent having an abnormally high number of friends and acquaintances within his sociability network who are listed in the register and who have not been named by another respondent within the sample) not mentioned elsewhere; the list mode which allows all the individuals named in the survey to be closer to the sampling frame by all members of the connected (networked) component of the target population.11. The Commission also observes that a gift voucher in the amount of 15 euros will be given at the end of the administration of the questionnaire to compensate

respondents for the time they have devoted to it, as well as a gift voucher of amount of 20 euros in the event of dissemination of information relating to the survey by the respondent to his social network. It recalls that the delivery of gift vouchers, which cannot be conditional on the collection of personal data, is only intended to compensate the respondent for the time spent and has no impact on the legal basis mobilized for the planned processing (execution of a mission of public interest).¹² Finally, the Commission notes that, as this is not a mandatory survey, the consent of the persons concerned will be requested orally by the surveyors before the start of the questionnaire. On the purposes of the processing¹³. The main objectives of the CHIPRE survey are to: remedy the lack of detailed quantitative data on Chinese people born in China and residing in France and thus build a quantitative database to enrich knowledge of Chinese people living in Paris and the surrounding areas Parisian ; measure the heterogeneity of this population; explore the links between Chinese populations of different geographical origins in China and non-Chinese populations, as well as the role of sociability networks in job search and adaptation to the host society; test the existence of a model of selective adaptation of the members of this population, particularly according to the region of origin in China; to study the determinants and perceptions of possible discrimination within and outside the community (within French society).¹⁴ The Commission, which takes note of the clarifications provided that the survey project pursues only scientific research purposes, considers that the purposes pursued by the planned processing are determined, explicit and legitimate, in accordance with the provisions of b) of 1. of Article 5 of the GDPR. On the data processed ¹⁵. Firstly, the Commission observes that the purposes pursued involve the collection of numerous personal data of a very varied nature from the persons surveyed, some of which fall into the category of personal data. say sensitive within the meaning of Article 9 of the GDPR, in particular data concerning health and data which reveal origins or religious beliefs.¹⁶ It recalls that, in decision no. 2007-557 DC of November 15, 2007, the Constitutional Council ruled that if the processing necessary for conducting studies on measuring the diversity of people's origins, discrimination and integration may relate to objective data, they cannot, without disregarding the principle set out in Article 1 of the Constitution, be based on ethnic origin or race .¹⁷ The Commission observes that the survey project, which mainly aims to collect objective data, is not based on the alleged racial or ethnic origin of the respondents. While some questions such as those in the sections on past and present employment and those on experiences of discrimination involve the collection of subjective data, she believes that this survey is not intended, even indirectly, to classify respondents depending either on their alleged declared ethnic or racial origin, or on an ethno-racial reference. Under these conditions, the Commission considers that the questions asked are not contrary to the aforementioned

decision of the Constitutional Council.¹⁸ Secondly, it notes that it is planned to collect data from respondents to the main survey relating to many people around them, both relatives (children for example) but also six people in their network sociability, in particular contact, so that INED can extend the survey sample and describe the nature of the first respondent's sociability network.¹⁹ While the Commission takes note of the reasons invoked in support of such collection, namely the absence of a national sampling base allowing the random selection of persons belonging to the target population, it considers, with regard to the aim pursued by the planned processing, that only the identity, contact and relationship data with the first respondent are necessary for the investigator to contact the person concerned directly within the maximum period of one month from the collection from the first respondent responding.²⁰ Having recalled these elements, the Commission considers that it will be up to INED to contact the member of the first responder's social network to inform him, including when the interviewer has to tell him that he is not eligible for sampling. In which case, the personal data concerning him must be deleted at the end of the communication.²¹ The Commission also notes that the respondent does not obtain the consent of the members of his social network to the processing of their data by INED, but only their agreement to be contacted by an investigator within the framework of the scientific methodology. It follows from these elements that such an agreement cannot be a consent of the members of its network to the processing of their personal data, within the meaning of the GDPR.²² Thirdly, the Commission notes that it is envisaged that the interviewer will indicate observations relating to the attitude of the respondent (hostile, impatient and agitated) and the acceptability of the register (how many contact details of his friends data; request justifications on the reason for requesting the register etc.) in a Quality Questionnaire . In view of the purposes pursued by the planned processing, it considers that such collection does not appear justified. On the rights of individuals²³ The Commission notes that information will be provided, in French or Chinese, to the persons concerned, in the consent form to participate in the survey, and publicly via a mini-website available in French and Chinese. It notes that the consent form mentions the possibility of refusing to answer certain questions.²⁴ The Commission recalls that the information provided must comply with the information provided for in Article 13 of the GDPR and, with regard to persons whose data is collected indirectly, with the provisions of Article 14 of the GDPR and Article 79 of law n° 78-17 of January 6, 1978 as amended. In particular, it requests that the aforementioned media be modified in order to mention the legal basis of the processing or the retention periods of the data.²⁴ In addition, these same media indicate that the responses will be anonymized and that no use of the data will make it possible to identify the persons concerned. In this regard, the Commission

recalls that according to Opinion 05/2014 of the Article 29 Data Protection Working Party on anonymisation techniques, anonymisation is an irreversible technical process ensuring that data does not can no longer allow the identification of persons whether by individualization, correlation or inference.²⁵ Insofar as it has not been established that an anonymization process will actually be implemented in the context of the planned processing, it requests that the aforementioned materials be modified on this point.²⁶ The other rights of individuals do not call for any particular observation. On the recipients of the data²⁷. The Commission, which takes note of INED's commitment that only the INED cloud will host the personal data collected, recalls that it is necessary to implement the security measures necessary for the processing of data by researchers located on the territory of the United States with regard to the choice of the data controller not to organize a transfer of the data outside the European Union.²⁸ The Commission requests, in accordance with Articles 13 and 14 of the GDPR, that all categories of recipients, in particular the scientific community via distribution networks, appear in the information delivered to the persons concerned. On retention periods²⁹. It follows from the clarifications provided by INED that several retention periods are likely to be applied. The Commission considers that the retention period cannot exceed five years from the last scientific publication, which INED indicates will take place no later than 2023. Safety measures³⁰. In general, the Commission notes that specific training will be delivered to investigators in order to make them aware of the applicable data protection regulations. Similarly, it notes that the data collection form will be completed directly on the target database in a remote connection, which is likely to constitute good practice.³¹ Furthermore, although the Commission notes that a commitment not to download the data is planned between INED and the members of the research team located in the United States, it considers – with regard to the all the elements brought to its attention – that additional security measures should be put in place in order to meet the security requirement provided for in Article 32 of the GDPR. The Commission recalls, in any case, that this obligation requires the updating of safety measures with regard to the regular reassessment of the risks. In this respect, the Commission recalls that specific attention should be paid to the reassessment of security measures as part of the regular update of the impact analysis. For the PresidentDeputy Vice-President Sophie LAMBREMON