

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations



2020年，中国克服新冠肺炎疫情影响，成功组织开展了第七次全国人口普查（以下简称“七普”）。与历次人口普查相比，这次普查在广泛征求部门、专家、学者和基层意见、深入开展各类试点的基础上，充分总结借鉴历次普查经验，依托中国飞速发展的网络基础设施和大数据、云计算技术，在普查方法、流程及数据处理策略等多个方面进行了革新，通过更加精细化的质量控制，普查漏登率仅为0.05%，圆满完成查准人口、摸清底数的目标，也为世界各国的人口普查，特别是在疫情条件下开展大型普查提供了有益参考。

In 2020, China successfully organized the Seventh National Population Census (hereinafter referred to as the "Seventh Census") despite challenging conditions under the COVID-19 pandemic. Many government agencies, experts, academics, and grassroots personnel have been consulted as part of the Seventh Census, and several pilot projects were implemented. Building on this and drawing on the experiences from previous censuses, the Seventh Census sought to innovate in many areas, such as census methods, procedures, and data processing strategies, by utilizing China's rapidly developing network infrastructure, big data, and cloud computing technologies. Through these innovations and enhanced quality control measures, the census recorded a net undercount rate of only 0.05%, thus successfully meeting its expected objective to collect quality demographic data of the population. China's experience of completing the Seventh Census provides a useful reference for other countries preparing to conduct censuses, especially in challenging conditions such as the context of the pandemic.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

### 一、迎难而上：普查方案的出炉

作为在新冠肺炎疫情背景下开展的重大国情国力调查，七普面临着前所未有的组织实施压力：普查对象数量众多，工作人员队伍庞大，多层次组织管理任务艰巨；人户分离、人口流动、一户多房现象普遍增多，普查工作量以及锁定普查对象难度大幅提升；人们的隐私保护意识不断增强，居民配合意愿明显降低；新冠肺炎疫情的反复不仅增加了基层防控压力，更为普查组织实施带来诸多不确定风险；基层工作人员由于素养差异对信息化还存在一定抵触；普查资料逐级处理上报的“传统模式”面临着经费投入大幅增加、数据处理效能较低等问题，改革创新迫在眉睫。

针对以往普查的老问题和当前形势下的新需求，中国的普查设计者们在总结历次普查经验的基础上，对全面采用电子化、信息化方式提高普查效率和质量的可行性进行了充分论证，认为迅猛发展的信息技术和快速扩大的用户群体能够为普查顺利开展提供坚实支撑。一方面，中国互联网络 and 智能终端设备的普及已经具备了规模基础，互联网普及率从2010年“六普”时的31.8%快速增长为

### 1. Rising to the challenge: formulation of the census programme

As a major survey that assesses China's national circumstances and capabilities in light of the COVID-19 pandemic, the Seventh Census faced unprecedented challenges in its organization and implementation: the number of census respondents was high, the size of the census workforce was significant, and the multi-level organization and management of the census was arduous; the phenomenon of people residing somewhere other than the address listed on their household registration, population migration, and the gradual increase in the number of housing units per household have greatly increased the workload associated with the census and the difficulties in contacting census respondents; the public's awareness of privacy protection has steadily increased, and residents' willingness to participate has significantly decreased; the various waves of the COVID-19 pandemic have not only increased the pressure on grassroots prevention and control, but also brought about uncertainties and risks to the implementation of the census; grassroots personnel (including supervisors and enumerators) had some resistance to information technology due to the uneven level of digital literacy; and conventional approaches of decentralized data processing and reporting faced challenges such as the significant increase in financial investment required and low data processing efficiency, which calls for imminent reform and innovation.

In view of the enduring challenges and experiences from previous censuses and the emerging demands, China's census designers looked into the feasibility of fully adopting an electronic approach based on information technology to improve census efficiency and quality. It was agreed that the rapid development of information technology and the growing number of Internet users could provide solid support for the census' smooth implementation. On the one hand, the

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

2020 年的 67.0%，移动网民数量从 2.8 亿猛增至 9.3 亿<sup>1</sup>，且城乡差异明显缩小，网民年龄结构趋向均衡。另一方面，云计算、大数据等计算机技术以及基础设施的发展能够为普查项目的上层和后端业务分别提供托管和技术底座；依托虚拟化、安全基座和现代通信等设施，能够实现数据生成和负载管理等服务，为普查业务的平稳运行提供更加便捷、方便实用的保障。

基于上述考虑，国家统计局早在 2017 年便明确了七普方案信息化、便捷化、实用化的设计理念，并本着“科学与可行、需要与可能、继承与创新”的原则，分别在甘肃、山东、广东、福建、湖北、四川和江苏 7 个省份开展了行政记录和手机信令数据应用、住房调查可行性、港澳台和外籍人员普查登记、现场比对核查、区域划分和地图绘制方法等 6 项国家级专项试点和 1 项综合试点。与以往普查试点最大的不同在于，七普的试点目的除了传统的优化普查方案、检验指标设置、积累普查经验外，更加注重方法的研究和可行路线的探讨。试点内容充分结合成熟的技

Internet and smart terminals have already reached a certain level of popularity, with the Internet penetration rate rapidly increasing from 31.8% during the Sixth Census in 2010 to 67.0% in 2020, and the number of mobile Internet users increasing from 280 million to 930 million during the same period.<sup>1</sup> Furthermore, the digital gap between urban and rural areas has narrowed significantly, and netizen age composition has become more balanced. On the other hand, the advancement of computer technologies and infrastructure such as cloud computing and big data can provide hosting and technical support for the top-level and back-end data of the census, respectively. Services such as data generation and load management can be realized using digital infrastructure including virtualization, security and modern communications, providing more efficient, convenient and practical guarantees for the census' smooth implementation.

Based on the foregoing, as early as 2017 the National Bureau of Statistics (NBS) stated that the design concept for the Seventh Census would prioritize informatization, user-friendliness and practicality. In accordance with the principles of "Science and feasibility, demand and capability, inheritance and innovation", six special pilots (including pilots on the application of administrative records and mobile phone big data, feasibility of housing survey, enumeration of residents from Hong Kong, Macao, Taiwan and from foreign countries, methods for cross-checking and verification, enumeration area (EA) demarcation and mapping) and one comprehensive pilot were implemented in seven provinces of Gansu, Shandong, Guangdong, Fujian, Hubei, Sichuan and Jiangsu. The Seventh Census pilots differed from previous census pilots in their emphasis on methodological research and deliberations on option feasibility, in addition to general

1. 第 46 次《中国互联网络发展状况统计报告》，中国互联网络信息中心 (CNNIC)。

1. *The 46th Statistical Report on Internet Development in China*, China Internet Network Information Center (CNNIC).

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

术路线和预想的改革目标，在顶层设计上贴近基层实操，视野更加广阔，过程更加务实。通过开展试点，中国的普查工作者在获悉新技术实践经验的基础上，对最初的普查设想进行了大刀阔斧的修订。例如通过试点验证了全面使用电子化采集的可行性，夯实了全面采用电子化方式开展普查的决心；探索了通过登记居民唯一身份标识进行比对、核查和应用行政记录、大数据的基本思路。

最终，在广泛征求相关部门、专家学者、基层一线工作人员意见的基础上，国家统计局把试点经验与来自各领域的意见进行了充分融合，制定了普查方案和工作流程（见图1），明确了普查内容、方法以及各环节的工作要求。从中国国情出发，本着少而精的原则，每一个指标的设置都从必要性、可获取性、实用性出发进行反复论证，体现严谨规范；每一个环节都经过充分讨论、征求基层意见和实践检验，确保科学可行。

七普的方案设计实现了中国历史上的多个首次：首次实现普查全流程电子化、信息化，各阶段工作实现精准对接；首次应用行政记录和企业大数据（包括网络运营商的位置大数据、电力大数据、通信大数据）辅助开展

optimization of the census programme, suitability testing of intended census questions, and accumulation of census experiences. Furthermore, the pilots' content fully integrated proven technical pathways and reflected the expected reform goals and the top-level design mirrored grassroots practices, resulting in a broader vision and more pragmatic census process. Through the implementation of the pilots, China's census workforce made major revisions to the original census plan based on the practical experiences with new techniques. The pilots, for example, validated the feasibility of electronic data collection, reinforcing the drive to conduct the census fully electronically. Furthermore, the basic idea of including the unique citizen identification (ID) number to enable cross-checking and verification with administrative records and big data was explored.

Finally, NBS fully integrated the pilot experiences with inputs from various fields, formulated the census programme and workflow (see Figure 1), and clarified the content, methods, and requirements of each component of the census through extensive solicitation of inputs from relevant government agencies, experts, academics, and grassroots personnel. Based on China's national context, and the "few but essential" principle, the development of each census question was fully justified for its necessity, availability and practicability to ensure rigor and standardization. Each component of the census was also thoroughly discussed, with grassroots input sought and practical testing carried out to ensure the census scientific and feasible.

The Seventh Census programme design achieved many firsts in Chinese history. The census was fully electronic and informatized for the first time, and each stage of work was seamlessly connected. This was the first attempt to use administrative records and big data from enterprises (including positioning data from telecommunications service providers, and big data from electricity meters and mobile phones) to support enumeration and cross-checking, which significantly

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

普查登记及核查，数据质量显著提升；首次利用互联网云技术、云服务在全国集中部署数据处理环境，对数据进行全流程安全管理和数据加密，避免各级重复投入并实现普查业务数据实时传输、快速处理；首次将公民身份号码纳入普查，运用数据比对复查技术，解决了以往人口漏报和重复计算难题；首次

improved data quality; and the first time that a centralized data processing environment was deployed across the country using Internet cloud technologies and cloud services to safely manage and encrypt data throughout the process, avoiding duplicated investment at all levels and achieving real-time sharing and rapid processing of census data. This was the first time that citizen ID numbers were recorded in the census, and data cross-checking and verification technique was used to address omission and duplication issues. For the first time, online self-enumeration via mobile Internet was made

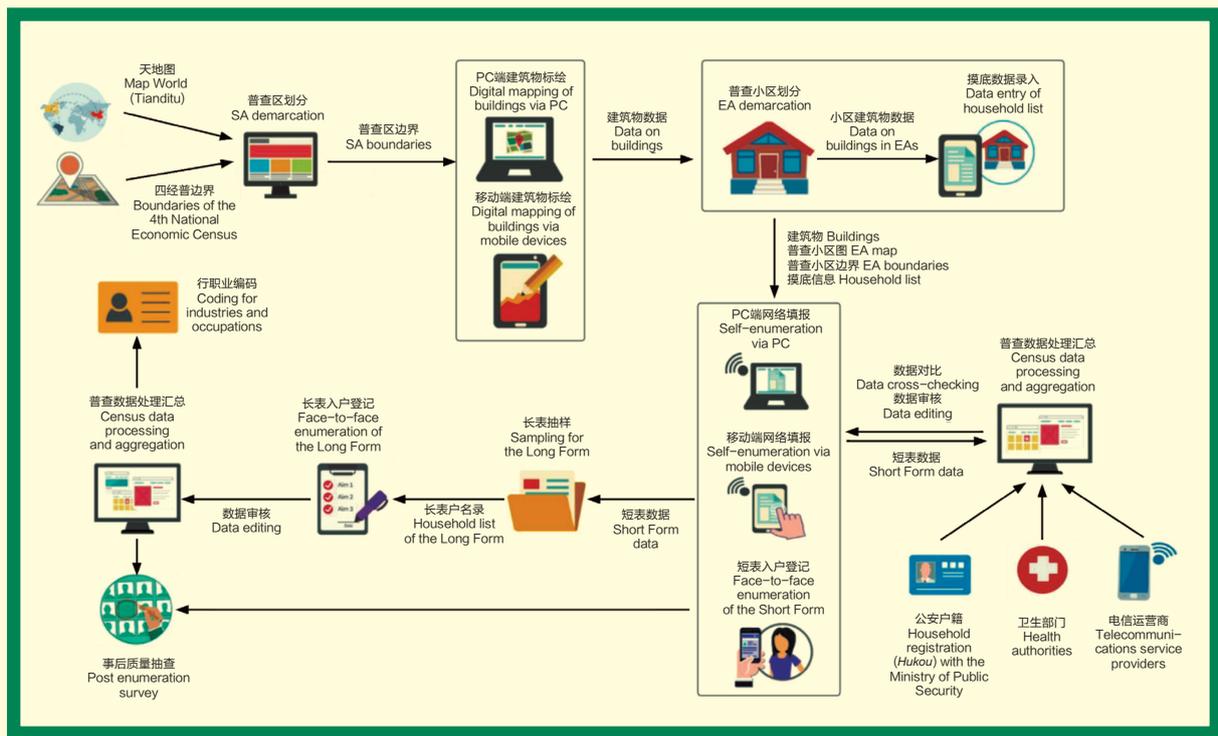


图 1 第七次全国人口普查业务流程图

Figure 1. Workflow of the Seventh National Population Census

注:

1. SA——普查区；EA——普查小区。
2. 第四次全国经济普查标准时点是 2018 年 12 月 31 日。
3. “天地图” 是国家测绘地理信息局建设的国家地理信息公共服务平台。

Note:

1. SA – Supervisory Area; EA – Enumeration Area.
2. The reference time of the Fourth National Economic Census is December 31st, 2018.
3. Map World (Tianditu) is the National Platform for Common Geospatial Information Services established by the National Administration of Surveying, Mapping and Geoinformation.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

部署移动互联网自主填报，充分适应居民个人隐私保护需求；首次依托互联网对普查队伍进行企业化管理和培训，节省大量管理费用的同时大幅提升了普查队伍的管理水平和工作能力；首次在重大疫情期间圆满完成普查任务，实现了人员无感染、工作无拖延、地域无遗漏。

### 二、扎实筹备：普查的事前工作

各级政府的高度重视是普查成功实施的根本保证。2019年10月31日，国务院印发《关于开展第七次全国人口普查的通知》（国发【2019】24号），提前一年对普查工作进行了部署，成立了由副总理韩正任组长、25个成员单位共同参与的国务院第七次全国人口普查领导小组，并在国家统计局设立领导小组办公室。地方各级政府按照通知要求，在普查期间组建了67.9万个普查机构，建立了覆盖省、市、县、乡、村的多层级普查网络体系，确保了普查工作有人抓、能落实，普查要求不变形、不走样。依托完善的普查网络，国家统计局在普查登记开始前组织开展了区域划分和绘图、户口整顿和行政记录资料整理、两员（普查员和普查指导员）选聘培训以及程序开发与测试等工作，为普查顺利开展奠定了基础。

available, catering to the personal privacy protection needs of respondents. It was also the first time that the census used the Internet to conduct enterprise-type management and training for the census enumerators and supervisors, saving significant management costs while vastly improving their management and work capacity. Finally, it was the first time that a census was completed successfully during a pandemic, with no COVID-19 infections among the census workforce, no delays in implementation, and no geographic areas omitted.

### 2. Vigorous preparation: key tasks before the census

High level attention from governments at all levels was fundamental for the successful implementation of the census. One year in advance, on October 31st, 2019, the State Council issued *the Notice on Conducting the Seventh National Population Census* ([2019] No. 24). The Leading Group of the State Council for the Seventh National Population Census was formed, led by Vice Premier Han Zheng and comprised of representatives from 25 ministries. In the NBS, a leading group office was established. In accordance with the requirements of the *Notice*, local governments established 679,000 census agencies at the provincial, prefecture, county, township and village levels, forming a multi-level census network to ensure that the census process was accountable and that the census requirements were followed as planned. Relying on the well-established census network and before the census enumeration began, the NBS conducted EA demarcation and mapping, verified and updated household registrations, reorganized administrative records, selected and trained enumerators and supervisors, and developed and tested the applications. This laid the foundation for the smooth implementation of the census.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

### (1) 区域划分和绘图

为明确普查区域界线，确保普查登记不重不漏、完整覆盖，国家统计局协调民政部门，提前半年暂缓实施各级行政区划调整，为普查区域划分提供了有利条件。信息化理念在区划绘图工作得到了充分体现，创新开发了“普查区划分与建筑物标绘系统”，将自然资源、规划、民政部门的地图和边界资料加载整合，借助遥感影像、天地图等相关地图资料进行实地勘察，利用云计算技术对数据进行处理，实现了全国各级实时在线制图、编辑、审核、上报，仅用时3个月便高效完成了全国600多万普查小区边界划分和上亿栋建筑物的数字化标绘工作。

### (2) 户口整顿和行政资料整理

七普继续采用在普查对象现住地和户籍地同时登记的方式。一方面可以最大限度发挥户籍行政记录的参考作用，另一方面也使得普查员更容易识别登记对象，便于培训和现场操作。为此，公安部门专门在普查开始前组织开展了户口整顿工作。通过户口整顿，规范了户籍记录的地域划分，清理了以往户籍记录中存在的重复、遗漏和应销未销的户口。在普查登记开始前，公安部与国家统计局联

### 1) EA demarcation and mapping

To facilitate EA demarcation, the Ministry of Civil Affairs froze administrative area boundaries six months before the census date, under the coordination of the NBS, so that no further jurisdictional changes were implemented. This was significant in clarifying EA boundaries as well as ensuring complete coverage of the population without omission or duplication. The EA demarcation and mapping fully reflected the concept of informatization. Furthermore, an “EA demarcation and buildings mapping system” was developed to integrate data on maps and boundaries from the departments of natural resources, planning, and civil affairs; field visits were conducted with support of mapping data such as remote sensing images and Map World (Tianditu); and data was processed using cloud computing. This enabled real-time mapping, editing, reviewing, and reporting across the country at all levels. It took only three months to complete the mapping of boundaries of more than 6 million EAs established nationwide, and hundreds of millions of buildings were labeled during the digital mapping process.

### 2) Verification and update of household registration (Hukou), and compilation of administrative records

The Seventh Census continued to enumerate people both where they were at the time of the census and where their household registration was. On the one hand, this increased the use of the administrative records of household registration as reference, while on the other hand, this made it easier for census enumerators to identify respondents. This also facilitated training and on-site operations. To that end, before the census began, public security departments specifically verified and updated household registration records. The geographical demarcation of household registration records was standardized through the verification and update of household registration, and issues of duplication, omission and invalid household registration records were resolved. The Ministry of Public Security and the NBS jointly

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

合下发通知，要求各级公安部门在做好保密工作的基础上，将户口整顿资料移交同级普查机构。除户口整顿资料外，各地普查机构还按照国家统计局的指导要求，积极协调卫生健康、民政、人社、医保等多种来源的行政记录，按照普查的行政区域归属进行整理，为现场登记工作打下良好基础。

### (3) “两员”选聘和培训

在人户分离现象普遍存在的社会环境下，采取现住地和户籍地两方面登记的普查方式，使得普查工作量大幅增加，对各级工作人员和普查员的数量需求远超以往。七普全国共选聘了超过 700 万名“两员”，并配备了一定比例的预备人员，其选聘、管理和培训面临极大考验。为此，国家统计局将普查队伍的建立作为普查准备阶段的重点工作，专门下发指导性文件明确普查员的胜任条件，并加大普查员管理、培训力度，借助“互联网+技术”打造了管理培训平台，对普查队伍进行企业化管理。所有普查工作人员的报名、注册、学习、考试、任职都采用线上和线下相融合的管理模式，在实时掌握了解队伍建设情况的同时，兼顾了普查员对普查信息化的适应过程。普查员在任职前都完成了考试认定，并在签署保密协议后，在线

issued a notice prior to the start of census enumeration requiring public security departments at all levels to share household registration information with census agencies while maintaining confidentiality. In addition, in accordance with the NBS guidance and requirements, local census agencies actively coordinated the sharing of administrative records from various sources, including health, civil affairs, human resources and social security, and medical insurance, and compiled the records by the census administrative areas. This laid a good foundation for on-site census enumeration.

### 3) Selection and training of enumerators and supervisors

As more people live outside their household registration address, enumerating people both at their place of current residence and their household registration have greatly increased the overall workload, and the size of the census workforce has grown significantly at all levels. The Seventh Census recruited over 7 million enumerators and supervisors across the country, and established a pool of reserve personnel. Their selection, management and training faced great challenges. To that end, the NBS considered the establishment of the census workforce as a key priority in the preparation stage, issued guidance to clarify enumerator qualifications, and strengthened their management and training. A management training platform was established using “Internet Plus Technology”, to implement an enterprise-type management of the census workforce. All census workers were screened, registered, trained, examined, and appointed using a management model that combined online and offline components. While this management model aided in informing real time recruitment progress, it also provided an opportunity for enumerators to adapt to informatization. Enumerators had to pass an exam and obtain certification before assuming the position. Enumerators were able to apply for certification online after signing the confidentiality agreement. In order to standardize the on-site work of

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

申领普查证件。为规范普查员现场工作，国家统计局还专门在数据采集系统中开发了普查员行为分析功能，强化对其工作流程的监督。

### （4）程序开发与测试

早在普查试点期间，国家统计局就委托第三方技术公司全程参与，借以深入了解普查工作流程，积累现场操作经验，据此制定了详细的用户需求，并在普查开始前进行公开招标，确定由国内实力较强、具有丰富互联网用户运营经验的软件公司负责普查系统(见图2)开发建设。产品功能同步交付各地测试使用，及时发现和解决问题，征集改进建议，充分结合顶层设计思想和基层实践经验，极大提升了开发效率。

普查程序的开发广泛借鉴了以往实践经验。在普查开始前，国家统计局先后组织了数次演习。通过软件功能测试、压力测试和安全性检验，结合工作需要强化了算力部署，调配了上千台数据处理服务器，确保普查数据在普查现场登记后可直接传送至国家统一的服务器集群进行初步处理和备份，真正实现了数据实时加密、实时推送、实时校验、实时共享。为确保数据处理时效性和安全性，专门建设了物理隔离的独立数据处理中心，负责普查核心业务处理和数据备份；普查数据采集端实现了自动保存、

enumerators, the NBS specially developed a behavior analysis function in the data collection system to strengthen supervision of enumerators.

### 4) Census application development and testing

As early as in the pilot period of the census, the NBS engaged a third-party technology company to participate in the entire process to gain an in-depth understanding of census processes and accumulate on-site operational experiences. As a result, detailed user requirements were developed, and a public bidding process was carried out to select and commission a strong domestic software company with extensive experience managing Internet users to develop the census application (see Figure 2). Product functions were delivered synchronously to various locations for testing. Issues and problems were quickly identified and resolved, and suggestions for improvement were collected taking into account both top-level design ideas and grassroots practical experiences. All of these increased the efficiency in developing the census application.

The development of the census application extensively drew on past experiences. Before the census began, NBS organized several rehearsals. Through the implementation of software functioning tests, stress tests and security inspections, computing power was strengthened based on the requirements of the work processes, and over a thousand data processing servers were deployed to ensure census data be directly transferred to the national designated server for preliminary data processing and backup after enumeration was completed on-site. This achieved simultaneous data encryption, feeding, cross-checking and sharing. An independent and physically isolated data processing center was built to ensure timely and secure census data processing and data backups. Automatic saving, reporting, and real-time editing were all possible with the census data collection terminals. The census workforce was required to confirm the assigned work tasks in the census application, which

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

报送以及实时审核；工作任务主动认领，既适应了普查员的工作习惯，又保证了数据安全，最大限度避免了重复工作；为应对偏远地区网络信号不好的问题，专门设置离线采集功能，并与线上采集数据实现互通校验；充分运用微

not only adapted to enumerators' work habits, but also ensured data security and avoided duplication to the extent possible. To address the problem of poor Internet connectivity in remote areas, an offline data collection function was developed, enabling interoperability and verification with online data collection. The census has fully taken advantage of the reach of WeChat, a mature social media app used in China. Online self-

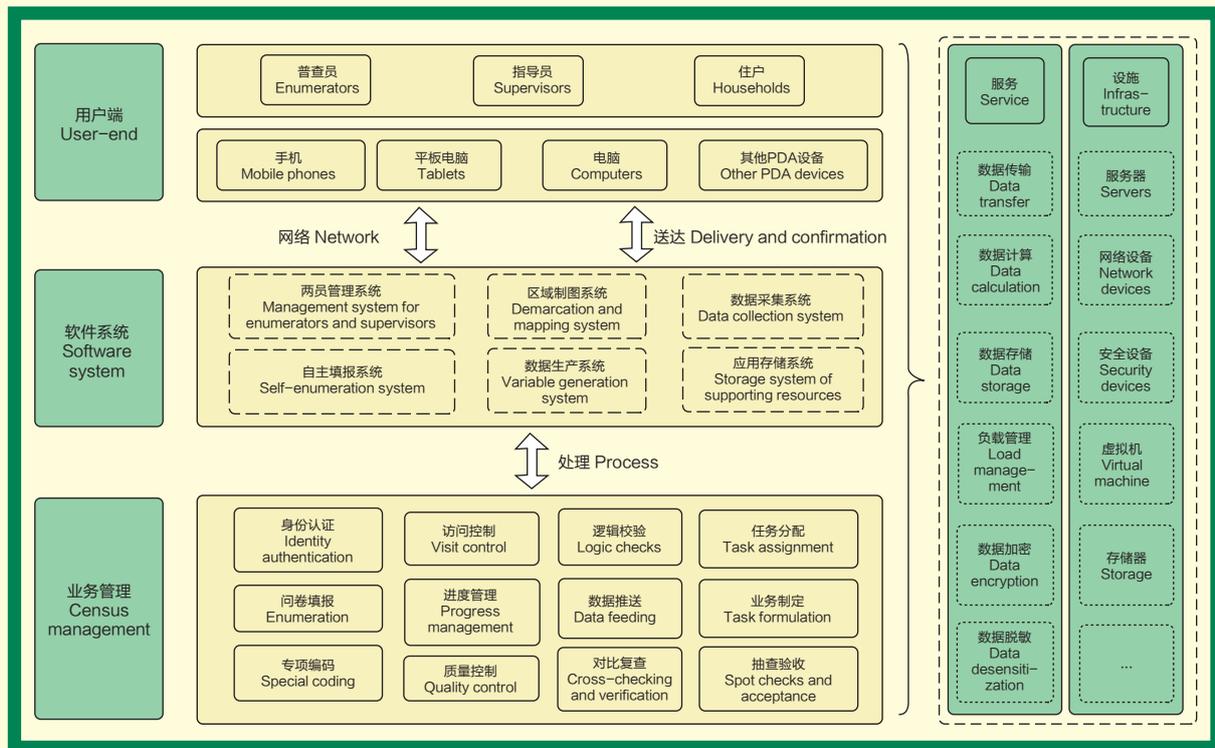


图 2 第七次全国人口普查信息系统处理架构及流程

Figure 2. Information processing framework and flow for the Seventh National Population Census

信这一成熟社交软件的触达能力，贴合用户常用操作习惯部署自主填报程序。普查登记期间，数据处理集群每天处理着数以亿计的普查信息，圆满完成了普查数据采集、质量控制和比对复查等核心任务。

enumeration was made available via a mini programme built in WeChat, designed based on user habits. During the census enumeration, the data processing cluster processed hundreds of millions of census records per day, completing the core tasks of data collection, quality control, cross-checking and verification successfully.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

### 三、精准实施：普查数据的采集和新技术应用

从2020年10月开始，七普正式进入数据采集阶段。数据采集大致可以分为摸底、登记和比对复查三个主要环节，质量控制和数据处理作为数据采集全流程必备措施始终贯穿其中。三个环节工作环环相扣，互为基础又互有交集，各种新方法、新技术的应用在这一阶段也得到集中体现。

#### （1）摸底：采集和建立普查登记基础

电子底册和指导住户自主填报是该阶段最明显的技术革新和亮点。摸底工作在普查登记时点前开展，目的是锁定各类普查对象。在这期间，普查员根据区划制图划分负责区域，结合电子影像图和户口整顿资料，对所辖区域内所有户籍登记地址和可能有人居住的建筑物、场所进行地毯式摸排。普查员的摸排结果通过移动通讯设备上传至国家统计局进行集中审核处理，形成全国统一规范的电子底册，并将居民家庭和建筑物以及行政区划的地理信息进行关联，为日后地理信息的建立打下必要基础。如果普查对象选择自主填报普查信息，普查员可向其提供账号密码或用于身份识别的二维码，并指

### 3. Sound implementation: collection of census data and application of new techniques

The Seventh Census officially began the data collection stage in October 2020. The three main components of data collection are: preliminary survey, enumeration, and cross-checking and verification. Quality control and data processing were required measures throughout the data collection. The three components' work was interlinked, and each component was based on and intersected with the others. Moreover, the application and use of various new methods and techniques were also evident at this stage.

#### 1) Preliminary survey: collecting data to establish the basis for census enumeration

The electronic household list and the guidance to households on self-enumeration were the most notable technical innovations during the preliminary survey. The preliminary survey was conducted prior to the census enumeration in order to confirm different census respondents. During this period, enumerators divided responsibilities according to the EA map. Referencing the materials on electronic map images and the verified and updated information on household registration, the census enumerators canvassed all registered household addresses and other buildings/places where people were likely to reside within responsible EAs. The enumerators' findings were uploaded to the NBS platforms via mobile devices for centralized review and processing, resulting in a nationally unified and standardized electronic list. This also linked the household to the building's and administrative division's geographic information, laying the groundwork for the future establishment of a geographic information system. If census respondents indicated a preference for online self-enumeration, enumerators provided an account with a password or a QR code for identity verification and guided them

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

导普查对象通过微信小程序完成自主填报。

大数据也在摸底期间尽其所长。在电力部门配合下，普查员可以通过手机 APP 快速且便捷地查询负责区域内各住房用电情况，根据 APP 提示的人口居住信息，辅助确定该住房是否空置，有效提高了摸底的准确性。通过摸底，各级普查机构在短时间内，对全国数亿栋建筑物进行了清查，明确了下阶段登记任务，并指导了 3000 多万住户进行自主填报。

### （2）登记：将公民身份号码纳入普查和新方法的应用

与历次普查相比，七普在指标设置方面最明显的变化是将公民身份号码纳入普查内容，为准确判断人口重漏报、应用行政记录和大数据辅助核查普查登记信息提供了基础，也为日后部门资料的共享应用奠定了良好基础。

七普在数据采集阶段也采用了其他国家常用的长短表登记技术。不同的是，七普的长短表登记不再同时开展。短表登记先期进行，主要目的是查清全国的人口数量和分布，指标设置相对少而精；短表登记结束之后，根据登记结果由国家平台统一抽取 10% 的住户进行长表登记，主要目的是查清人口的各种结构，指标的

through the self-enumeration process using the mini programme built into the WeChat app.

The use of big data was another highlight during this period. With support from the State Grid, enumerators were able to check the electricity consumption of each household in the responsible EAs easily and efficiently through a mobile application. This assisted in determining whether a residence was vacant based on information provided by the mobile application, effectively improving the accuracy of the household listing data. Finally, census agencies at all levels canvassed hundreds of millions of buildings across the country in a short period, defined the enumeration tasks for the next stage, and guided over 30 million households through the self-enumeration process.

### 2) Enumeration: including citizens' ID numbers in census and using new methods

Compared with previous censuses, the most noticeable change in the Seventh Census in terms of census questionnaire design was the inclusion of citizens' ID numbers. This allowed for the accurate detection of duplicates and omissions of population and the verification of census data through the use of administrative records and big data. It also laid a solid foundation for the sharing and use of data between sectoral departments in the future.

The Seventh Census also used both the Short Form and the Long Form, as commonly used in other countries during the data collection stage. However, the Seventh Census stopped the practice of enumeration using the Short Form and the Long Form at the same time. The Short Form was used first to determine the size and distribution of the population across the country, with limited yet carefully selected census questions. Once the Short Form was completed, 10% of the households enumerated were centrally selected by the national platform for the Long Form. The main purpose of the Long Form was to determine the demographic

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

设置更加详细全面。这种先登记再统一抽样的方法彻底解决了以往由于缺乏完备抽样框以及基层人为选择而导致的样本偏差。

信息技术的应用在工作衔接和提升数据处理效能方面发挥重要作用。如摸底期间形成的电子底册，在普查时点前统一通过身份认证系统分配至每名普查员，普查员使用移动终端登陆即可自动领取工作任务；数据采集程序内置公民身份号码校验规则以及上千条指标间、人间、表间的逻辑审核规则，数据上报同时进行无感化审核，错误清单自动定位，未完成工作智能提示；短表登记数据实时形成长表抽样框，“一键”完成全国长表样本的抽取和已登记内容的推送；采用大数据技术分析普查员工作行为信息，监测普查进度和异常行为，严防可能出现的人为失误；各工作流程均采用了线上抽查验收等方式。上述新技术的应用不仅使登记期间的各种业务场景设计成为现实，而且有效地提升了普查登记的准确性和时效性。

### （3）比对复查：普查技术的重大创新

七普最明显、最关键、最具影响的变化是比对复查的工作安排和技术应用。普查方案设计之初，国家统计局就将探索一种能够有效解决数据重漏报的方法作为首要任务，在认真研究以

compositions of the population, and hence the relevant census questions were more detailed and comprehensive. This approach of enumerating first for the Short Form and then selecting sample households for the Long Form solved the issue of sampling bias observed in the past caused by the arbitrary selection of households at the local level due to lack of complete sampling frame.

The use of information technology played an important role in work integration and improvement of data processing efficiency. For example, the electronic household list was uniformly assigned to each enumerator via the identity authentication system before the census began, with enumerators automatically receiving work tasks by logging into the mobile terminal. The data collection process includes built-in citizen ID number validation rules, as well as over a thousand logical rules on census questions, persons and forms for data editing, where automatic review was performed while data was submitted, errors were flagged, and reminders were created for incomplete work tasks. The Short Form data formed the sampling frame of the Long Form in real time and “one-click” was all that was required to select the national sample for the Long Form and transfer the collected information. Big data was used to analyze enumerators' behaviors, monitor census progress, identify anomalies, and prevent human errors. Each of the work processes also used approaches such as online spot-checks and data acceptance. The use of these new techniques not only operationalized the census scenario design during enumeration, but also effectively improved census enumeration accuracy and timeliness.

### 3) Cross-checking and verification: a major innovation in census techniques

The most evident, critical and influential change to the Seventh Census is the application of techniques on cross-checking and verification. During the early

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

往查重补漏的困难因素并有针对性地组织开展大规模试点后，形成了第七次全国人口普查数据比对复查的工作思路：通过普查数据和行政记录数据的外部比对，确定漏报人口；通过普查登记数据之间的内部比对，确定重报和错报人口；结合大数据技术手段建立全新的数据排重和补漏工作方法，在全面提高数据质量的前提下最大限度减少现场核查工作量。

基于以上设想，在设计比对复查具体流程时，国家统计局确定了比对核查 4 项基本原则：一是要确保每一公民身份号码都被核实后登记；二是所有公民身份号码对应的经核实登记人口都要确定为某地常住人口；三是所有重复登记的公民身份号码所对应的人口，只能确定为一个地区的常住人口；四是新出生没有公民身份号码的人口，以登记信息为准，确定其常住地。根据这 4 项基本原则，配套制定了行政记录比对和数据排重确定常住地的基本流程。

首先，通过将人口普查登记信息与行政记录比对，对人口的漏报和错报进行筛查、补登和改正（见图 3）。

其次，依据公民身份号码对普查登记信息进行排重，确定有效信息，并据此判断唯一常住地（见图 4）。

design of the census programme, the NBS prioritized the exploration of a method that can effectively resolve the challenges of duplication and omission. After careful analysis of the difficulties in previous censuses, and the organization of large-scale and targeted tests, the strategy on cross-checking and verification for the Seventh Census was developed: specifically, omitted populations were identified through external comparison of census data with administrative records, while duplicated and incorrect records were detected through internal comparison of census data; and a new method was established through the use of big data to eliminate duplication and avoid omission, minimizing the workload of on-site verification and ensuring the improvement of overall data quality.

The NBS specified four basic principles while designing the specific processes for cross-checking and verification based on the above strategy: first, ensure all citizen ID numbers are verified and recorded; second, a verified citizen ID number must correspond to a place of usual residence; third, all duplicated citizen ID numbers should be cleaned to correspond to only one place of usual residence; fourth, for newborns without citizen ID numbers, their place of usual residence should be determined based on the census enumeration information. The basic procedures for determining the place of usual residence were developed based on these four principles by cross-checking administrative records and eliminating duplicates.

First, through cross-checking census enumeration information with administrative records, omissions and errors were identified, retro-enumerated and corrected (see Figure 3).

Second, census enumeration information was cross-checked against the citizen ID numbers to eliminate duplicates and verify the validity of the information, based on which the single place of usual residence was confirmed (see Figure 4).

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

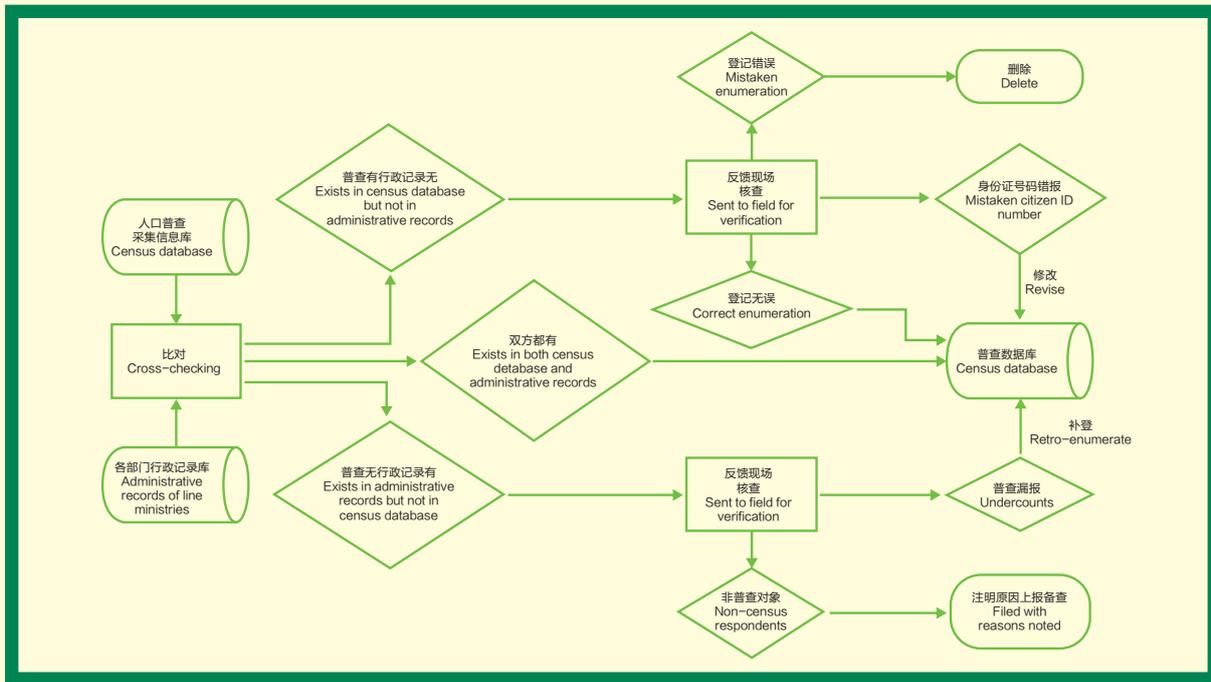


图3 第七次全国人口普查行政记录比对流程

Figure 3. Flowchart on cross-checking census data with administrative records

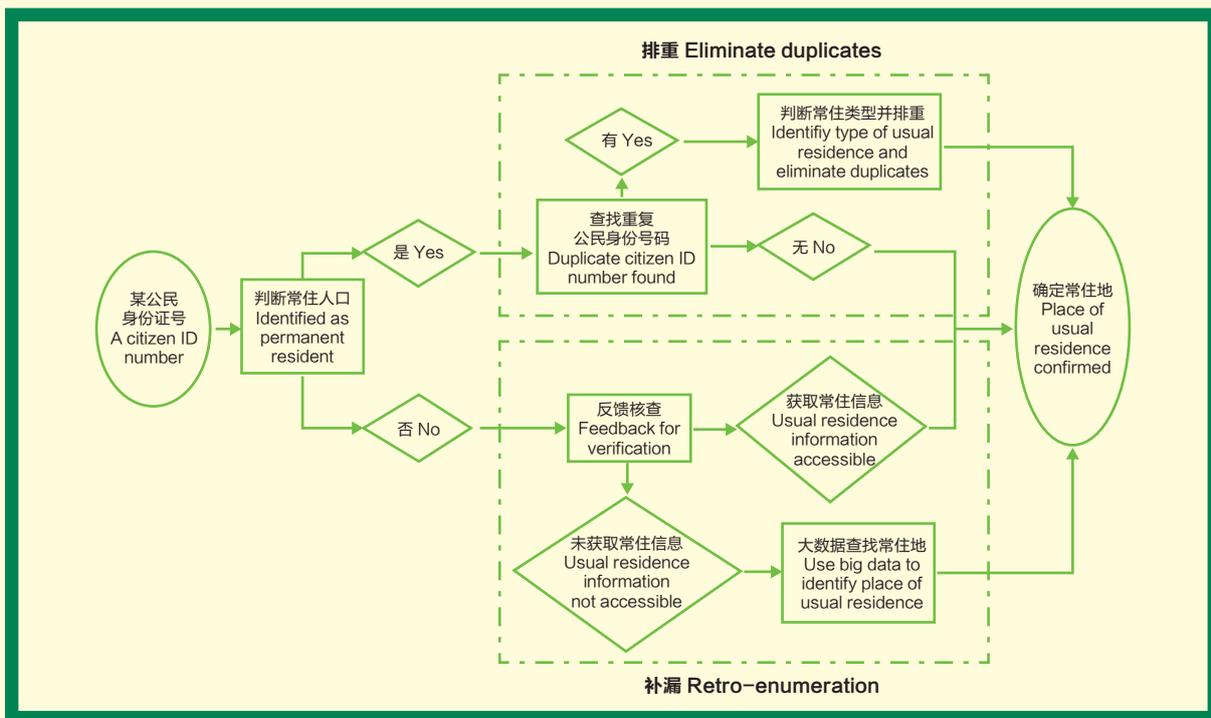


图4 第七次全国人口普查排重确定常住地流程

Figure 4. Flowchart on the elimination of duplicates and identification of place of usual residence

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

新流程的设计对数据处理能力及其算法优化提出了较高要求，每个人的登记信息都要和全国 14 亿人的登记信息进行比较的同时，还要和来自其他部门的行政记录进行比对生成核查记录，且时效性要求高。为此国家统计局又设计了一套全新的数据处理算法，以抽丝剥茧的逐步处理方式，将全部登记信息分别在现住地和户籍地进行编码分类，甄别无效记录的同时将问题信息直接通过国家平台向基层普查机构反馈核查。对个别无法核实的数据，通过通讯大数据分析手段，辅助判断流动人口的常住地。

在周密的事前部署和强大算力支持下，短表登记结束后，国家统计局仅用时 3 天便完成全部普查短表登记数据内部以及与公安、卫健行政记录的比对，对发现的问题按照地区和类型进行了详细分类，第一时间推送至普查一线工作人员。各地工作人员对照核查清单，统筹安排长表登记和比对复查工作，逐户上门核查，于登记结束一个月内便完成了核实登记，有效提升了普查登记质量。大数据分析技术在这个环节再次锦上添花，一是利用通讯和网络运营商的位置数据，对各地区活动设备数量进行监测，模拟常住人口分布，与普查登记数据进行比较，

The enumeration information of each individual needed to be not only compared with that of 1.4 billion people across the country, but also cross-checked against administrative records from other departments to generate verification records, on top of which there were strict requirements on timeliness. Such a new design of the work processes posed high demands for data processing and algorithm optimization. As a result, the NBS created a new set of data processing algorithms that coded and classified all enumeration information by the place of current residence and household registration in a step-by-step and meticulous manner. Once invalid records were identified, the national platform directly fed them back to the grassroots census agencies. The analysis of big data on telecommunications assisted in determining the place of usual residence for migrant populations for certain unverifiable data.

After completing the Short Form enumeration, the NBS completed the internal comparison of the Short Form data and the external comparison with the public security and health administrative records in only three days, thanks to meticulous preparation and the support of powerful computing capacity. The identified errors were classified in detail by region and type, and shared with front-line enumerators at the earliest time. Enumerators from across the country subsequently reviewed the checklist, conducted visits door-to-door for cross-checking and verification, as well as Long Form enumeration. All the work was completed in one month, and the quality of census registration was effectively improved. Big data analysis technology has once again risen to prominence in this process. First, positioning data from telecommunications and network operators was used to monitor the number of active devices in each region, simulate the distribution of permanent residents, and assess accuracy of census enumeration data. Second, positioning data was gathered in order to pinpoint the population whose current residence

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

判断人口登记的准确性。二是对普查未确定现居住地的人口进行定位统计，辅助各地进行查疑补漏。实践证明，比对复查的实施不仅有效解决了以往普查人口重报、漏报的痼疾，而且将部门行政记录、大数据应用与传统普查技术进行了整合，既将以往的理论研究变为现实，又给未来的人口统计监测提供了重要参考。

### 四、圆满收官：事后质量抽查和数据发布

事后质量抽查结果显示，七普漏登率仅为0.05%。七普对数据质量的高标准要求超过以往任何一次普查，不仅制定了全流程质量控制措施，而且要求各阶段必须逐级进行质量验收，相关的文档资料都要进行存档。在数据采集阶段结束后，还按照七普总体安排专门组织开展了普查事后质量抽查。全国共抽取141个县，406个普查小区，约3.2万户、10万人，采取双随机（随机抽取检查区域、随机选派检查人员）、异地交叉检查，并招聘社会监督员跟随监督，全程公开透明。

大规模的数据处理同步进行。由于采用了电子化登记方式，七普不需要再安排数据录入和编码工作，行业、职业的专项编码采用了智能编

was not determined in the census, and to assist local areas conducting verification and filling out omitted records. Cross-checking and verification have been shown in practice to not only effectively resolve the long-standing challenges of duplication and omission in previous censuses, but also integrate administrative records from government departments, big data technology, and conventional census techniques. The innovations of the Seventh Census made previous theoretical research a reality and provided an important reference for future population surveys and monitoring.

### 4. Successful completion: post enumeration survey and data publication

The results of the post enumeration survey revealed that the net undercount rate of the Seventh Census was only 0.05%. The standard for data quality in the Seventh Census was higher than any previous census. Quality control measures were developed for the entire census process, data acceptance with quality was required at every level and at each stage of the census, and all relevant documents were properly archived. Following the data collection phase, the post enumeration survey was conducted in accordance with the Seventh Census' overall plan. The post enumeration survey was conducted in 141 counties and 406 EAs across the country with approximately 32,000 households and 100,000 people selected. A double random sampling method (selecting survey areas and interviewers randomly) was used, cross-checking was conducted by interviewers from another place, and observers were recruited from the public to monitor the entire process in an open and transparent manner.

Large-scale data processing was performed synchronously. The use of electronic enumeration meant that data entry and coding was no longer necessary, and

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

码 + 人工的方式进行，减少源头性差错的同时节省了大量时间。与此同时增加普查数据的预处理，在国家层面部署数据处理环境，对全部普查数据进行再次审核，并将业务数据（指普查采集的关于人和住户的数据）与过程数据（指普查采集过程中出于工作需要由采集程序独立生成的数据）进行区分和备份。普查采集的所有信息在这个阶段都要进行重新审视和衍生计算，并依据计算结果，对数据进行再次分类，最终形成全国常住人口、死亡人口、港澳台和外籍人口数据库，以便在此基础上汇总计算人口总量及各种人口结构数据。数据汇总、分析解读、公报起草等工作多管齐下，2021年5月11日，国务院人口普查办公室、国家统计局召开新闻发布会，介绍第七次全国人口普查主要数据结果并回答记者提问。同时，国家统计局官网发布第七次全国人口普查第一号至第八号公报，截至2021年6月中旬，全国31个省（区、市）均按照统一部署和要求陆续发布了本地区普查公报，标志着第七次全国人口普查圆满完成。

the special coding for industries and occupations was done through the combination of "computer + manual coding", which significantly reduced source errors and processing time. Concurrently, pre-processing of census data was conducted where a data processing system was established at the national level and all census data was verified once more. Statistical data (refers to the data on people and households collected by the census) and para data (refers to the data separately generated by the census process) were differentiated and backed up. At this stage, all of the data collected by the census was reexamined, calculated, and reclassified accordingly. As a result, the national databases on the permanent residents, deceased population, as well as residents from Hong Kong, Macao, Taiwan and from foreign countries were compiled, which served as the basis of the aggregation and calculation for total populations and various demographic compositions. Data were published through various ways, including tabulations, analysis and interpretations, and communiqués. On May 11th of 2021, the Leading Group of the State Council for the Seventh National Population Census and the NBS held a press conference to present the main data of the Seventh Census and answered questions from reporters. Simultaneously, the NBS published eight communiqués of the Seventh Census (No.1 to No.8) through its official website. As of mid-June 2021, all 31 provinces, autonomous regions, and municipalities across the country successively issued respective local census communiqués in accordance with the overall arrangement, marking the successful completion of the Seventh Census.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

### 五、继往开来：普查的经验和反思

在普查方案设计方面，科学严谨的工作态度和坚持改革创新是普查成功的关键因素。从试点到实施，中国普查设计者始终锚定求真务实目标，广开言路、深入研究，切实加强顶层设计，进行了大胆创新和积极探索，普查各阶段工作融入了更多科技元素，实现了更多方法和技术的突破，极大地提高了普查工作质量和效率。其中，公民身份号码的采集以及围绕其开展的审核、比对和质量控制、行政记录和大数据应用等一系列普查新方法、新流程的设置，是这次普查最鲜明的革新措施，也是普查成功的重要保证，更为日后人口统计监测制度的优化提供了宝贵的实践经验。

普查之后再看普查，仍有一些问题需要进一步反思。首先，行政记录共享渠道有待进一步畅通。行政记录在七普中的应用仍局限在查清人口数量方面，受制于各部门数据保密要求，数据交换过程繁琐且内容偏少，难以对人口结构和社会属性进行更多补充；普查之后，由于多部门共享的制度及渠道尚未建立，更广泛的应用及普查资料的深度开发仍然受限。

### 5. Forging ahead: reflections from the census experiences

In terms of designing the census programme, the scientific and meticulous work attitude as well as the persistence on reform and innovation were key factors in the census' success. From pilot to implementation, China's census designers always upheld the principles of truth, pragmatism, openness, and in-depth research, strengthened the census' top-level design, encouraged bold innovation and active exploration, incorporated many scientific and technological elements in each stage of the census and achieved breakthroughs in new methods and techniques, greatly improving the census' quality and efficiency. The most notable innovations of the Seventh Census were the inclusion of citizen ID numbers and the associated data editing, cross-checking, and quality control, as well as the establishment of new census methods and processes, including the use of administrative records and big data. This also provided an important guarantee for the census' success and valuable practical experience for future optimization of the population statistical monitoring system.

Looking back at the Seventh Census, certain issues require further reflections. First, the channels for sharing administrative records need to be further streamlined and expanded. In the Seventh Census, administrative records were only used to determine population size. Subject to the data security requirements of various departments, the data sharing process was cumbersome, and the content was limited, making it difficult to further elaborate on demographic compositions and social characteristics of the population. After the Seventh Census was completed, as the mechanism and channels for multi-sectoral data sharing was not yet established, wider application and in-depth analysis of census data were still constrained.

# 中国人口普查：实践与创新

## China's Population Census: Experiences and Innovations

其次，普查筹备时间需留有余量。七普事前准备阶段总体安排与六普基本一致，但由于采用了大量新技术手段，缺少可借鉴经验，基础设施的采购建设与软件测试、调整、开发耗费较长时间，为保证普查如期、顺利开展，个别新技术未能投入使用，例如在数据采集端建立电子地图辅助普查员开展摸底、建立统一规范的建筑物和住户电子地址、开发独立的在线测试系统、通过第三方软件逆向查询本户地址、登记期间追加式进行数据比对等技术设想，因技术相对复杂，耗资较多，开发时间较长等原因最终未上线使用。

最后，新技术的使用并不意味着舍弃传统普查技术。普查各阶段虽然应用了大量新技术，但是传统普查技术（如手工快速汇总、纸介质表格登记、光电扫描识别和网页录入等）仍然作为普查的预案，在应对极端天气、偏远地区基础设施建设薄弱、疫情防控等客观因素带来的挑战时发挥了关键作用，是普查成功开展的重要保证。

Second, there should be more time reserved for preparations for the census. The overall plan of the preparatory stage of the Seventh Census was essentially the same as the Sixth Census. However, due to the adoption of many new techniques and the lack of referential experience, the procurement and development of infrastructure, as well as software testing, adjustment and development took a long time in the Seventh Census. Some new techniques were not fully implemented in order to ensure the census ran smoothly and on time. For example, due to the complexity of the process, several technical ideas were not implemented, such as developing an electronic map at the data collection terminal to assist enumerators in conducting the preliminary survey, developing a unified and standardized electronic address list of buildings and households, developing an independent online testing system, reverse-searching household addresses using third-party software, and conducting additive data cross-checking during enumeration.

Finally, the use of new techniques did not rule out conventional census techniques. Although a large number of new techniques were applied at each stage of the census, conventional census techniques (e.g., quick manual tabulations, paper form enumeration, optical character recognition, and web page entry) were still included in the census programme, and played a key role in coping with the challenges brought by extreme weather, weak infrastructure in remote areas, and pandemic prevention and control measures. The conventional census techniques were an important guarantee for the successful implementation of the census.