

ITEM FOR FINANCE COMMITTEE

CAPITAL WORKS RESERVE FUND

HEAD 710 - COMPUTERISATION

Census and Statistics Department

New Subhead “Information Technology Equipment and Services for the 2021 Population Census”

Members are invited to approve a new commitment of \$202,680,000 for acquiring information technology equipment and services for the 2021 Population Census.

PROBLEM

The next population census will be conducted in 2021. Given the sizable scale and complexity of the operation, it would not be practically possible to conduct the 2021 Population Census (21C) without adequate information technology (IT) support.

PROPOSAL

2. The Commissioner for Census and Statistics, with the support of the Secretary for Financial Services and the Treasury and the Government Chief Information Officer, proposes to create a new commitment of \$202,680,000 for acquiring IT equipment and services for the 21C.

JUSTIFICATION

IT Support for the 21C

3. The 21C is a large-scale and complex operation involving the enumeration of around 2.7 million households during the 36-day data collection period (tentatively scheduled for 30 June to 4 August 2021); the recruitment and

/training

training of around 18 000 temporary field workers (TFWs); the processing of completed questionnaires within a short period of time; and the dissemination of results in batches from February 2022. In designing the 21C operation, the prevalence of households living in subdivided units and industrial buildings and households with ethnic minorities in different geographical areas will be taken into account. Overall speaking, about nine-tenths of the households in Hong Kong will be subject to simple enumeration to provide basic demographic information using a “Short Form” questionnaire, and one-tenth of the households will be subject to detailed enquiry on a broad range of demographic and socio-economic characteristics by a “Long Form” questionnaire.

4. As the results of population censuses/by-censuses are vital to the Government for planning and policy formulation and are important to the private sector for business and research purposes, the data topics should reflect the major concerns of society at the time. In view of public concerns on social issues such as housing, labour supply and ethnic minorities in recent years, new data topics on (i) subdivided units; (ii) floor areas; (iii) working hours; and (iv) ability to read/write languages were introduced in the 2016 Population By-census (16BC) and will continue to be included in the 21C. To address the imminent challenge of population ageing, it is proposed that an additional data topic on “elderly persons requiring care” will be introduced in the 21C to gauge the elderly’s demand for care and support services and obtain basic information on their carers. Further research and engagement with stakeholders will be carried out before the data topics of the 21C are finalised.

5. Progressing with the times, IT will be used to enhance the experience of respondents in providing the required census data. For instance, there have been various changes in the modes of enumeration during the past decade. To suit the lifestyle of Hong Kong people, e-Questionnaires (eQ) were first introduced in the 2011 Population Census (11C) for desktop computers. The eQ was made mobile-friendly in the 16BC and will continue to be used in the 21C such that respondents, particularly the younger generation, can provide data online using smartphones or mobile tablets anytime and anywhere. Mobile tablets, which were first used by enumerators in face-to-face interviews in the 16BC, will continue to be used in the 21C to replace traditional paper questionnaires. The use of mobile tablets can enhance both user experience and data quality as instant data verification can be made such that the respondents will not be visited again for clarification and follow-up. Telephone interviews were first introduced in the 16BC, which provide more convenience to respondents, especially the elderly, who may require the guidance of TFWs in providing the data. To enable more efficient telephone interviews to be conducted, a dedicated computer sub-system will be developed in the 21C for the purpose.

6. A two-phase multi-modal data collection approach will be adopted in the 21C. In the first phase of the data collection period (tentatively from 30 June to 17 July 2021), eQ, telephone interviews and postal returns with pre-paid envelopes (for “Short Form” only) will be used. These data collection channels are more cost-effective and will be heavily promoted in the publicity campaign. Face-to-face interviews will be introduced as an additional channel during the second phase (tentatively from 18 July to 4 August 2021). For households which have not provided data through any of the above channels, the trained TFWs equipped with tablets will visit and conduct face-to-face interviews with them. An integrated IT system is needed to support the multi-modal data collection as well as the subsequent data processing and dissemination work.

7. In line with international trend, the Census and Statistics Department (C&SD) will make use of more administrative data in the 21C to reduce respondent burden and enhance data quality.

8. In 2018, the C&SD commissioned a feasibility study to examine the technical requirements and financial implications of the proposed IT system. A Business Process Re-engineering (BPR) study was also conducted to identify opportunities for streamlining and improving the workflow of the 21C. In view of the sizable scale and complexity of the census operation and with reference to the experience in past population censuses/by-censuses, the support of a tailored IT system is essential for conducting the exercise in a cost-effective way.

Proposed IT System for the 21C

9. The feasibility cum BPR study, completed in November 2018, proposes to enhance and reuse the IT system developed for the 16BC as far as practicable. The study also recognises the need for enhancement and new functionalities to meet the changing requirements of the 21C, in particular upgrading of the system to support the sizable-scale operation and provision of better multi-modal data collection services. It therefore recommends that the IT system for the 21C should be established using a combination approach involving enhancement and consolidation of the sub-systems of the 16BC into 15 sub-systems and development of one new sub-system for conducting telephone interviews. This approach represents the most cost-effective solution in developing the IT system for supporting the pre-census preparation, data collection, imputation and estimation, data dissemination and staff management in the 21C. As for the IT equipment used in the 16BC, most of them have either been re-deployed for use in other IT projects of C&SD, donated to non-governmental organisations as planned, or are due for replacement.

/New

New Functions and Enhanced Features

10. To provide better services to the public, and to improve the efficiency and quality of the 21C operation, the following new functions/enhanced features will be implemented in the IT system –

(a) *Enhancing system capacity and scope*

The system capacity, in terms of hardware, software and communication network, will be enhanced to support the much larger scale of the 21C as compared with the 16BC. The 16BC only involved a detailed enquiry by “Long Form” on about one-tenth of households in Hong Kong, whereas all households in Hong Kong will be enumerated in the 21C (about one-tenth using the “Long Form” and the remaining nine-tenths using the “Short Form”). Compared with the 11C, it is projected that the total number of households will grow by 12%, thereby further raising the system requirements. Moreover, the scope of most of the system components implemented in the 16BC will be enhanced to cater for the need to process two different questionnaire forms for two groups of respondents.

(b) *Enhancing multi-modal data collection services*

The 21C will offer multiple modes of data collection, including self-enumeration by eQ or by postal return (for “Short Form” only), computer-assisted telephone interviews, and face-to-face interviews making use of mobile tablets. While respondents will be encouraged to adopt self-enumeration as far as possible, telephone interviews, which are more cost-efficient than face-to-face interviews, will be promoted to those respondents who do not opt for self-enumeration. A new sub-system for supporting telephone interviews will be developed to enhance the enumeration efficiency and provide better respondent experience.

(c) *Better fieldwork progress monitoring*

The assignment and fieldwork management functions will be enhanced to enable generation of instant and comprehensive progress, such as by mode of data collection, field centre and individual TFWs, as well as more flexible assignment allocation. As the enumeration progress by different modes can be monitored and analysed more closely, field assignments can be re-allocated in a more timely and convenient way to facilitate better utilisation of field resources.

/(d)

(d) *Enhancing data quality*

Advanced computer-assisted data validation features will be implemented in the telephone interview sub-system and mobile tablets for face-to-face interviews to facilitate early detection of reporting errors and clarification with the respondents instantly. This can help save the cost and time for re-contacting the respondents, reduce respondent burden, as well as improve data quality.

(e) *Increasing efficiency in statistical imputation and estimation*

The efficiency in statistical imputation and estimation in the 21C will be increased by migrating the process from the workstation platform to high-speed server platform. The estimation quality will also be enhanced through adopting more advanced statistical algorithms.

(f) *Facilitating wider use of 21C statistics*

To promote wider use of the 21C statistics by the general public, the user interface and search function of the Interactive Data Dissemination Sub-system (IDDS) will be enhanced to assist the users in locating, tabulating and charting the required statistics more easily. Infographics and more advanced interactive visualisation will also be introduced to facilitate understanding and interpretation of the 21C results by general users. Furthermore, in support of the Government's open data policy, Application Programming Interface (API) will be developed to enable sophisticated users, such as data analysts, researchers and programmers, in performing machine to machine query and eliminating replicated steps in accessing the census statistics.

(g) *Strengthening of IT security*

Strengthened security measures, such as Distributed Denial of Service (DDoS) protection, Web Application Firewall, Network Intrusion Prevention System and Mobile Device Management (MDM), will be implemented to protect the proposed IT system and its data contents from cyber-attacks. Data encryption to protect the collected data and end-point security protection to workstations will also be applied. The IT system will adopt security standards that are internationally recognised, such as Public Key Infrastructure for digital signing, and Advanced Encryption Standard and Transport Layer Security for data encryption, etc.

Privacy Data Protection and System Security

11. C&SD will conduct Privacy Impact Assessments at critical stages of the implementation of the IT system for the 21C, including system analysis and design stage, and prior to system rollout in order to ensure that the data protection principles promulgated in the Personal Data (Privacy) Ordinance (Cap. 486) are observed and complied with. In addition, C&SD will also engage independent auditors to conduct IT security risk assessment and security audits to ensure the effectiveness of those security measures in protecting information in the system. We will consult the Office of the Privacy Commissioner for Personal Data to ensure that the operation of the 21C is in line with the relevant requirements.

Anticipated Benefits

12. The proposed system will be critical to the smooth conduct of the 21C and hence the production of good quality statistics for use by a large number of users in both the public and private sectors.

13. The proposed IT system is designed to meet all the operational, timing, security and data quality requirements of the 21C. It re-uses the 16BC sub-systems with necessary enhancement and consolidation as explained in paragraph 9 above to maximise cost-effectiveness.

14. Apart from the benefits arising from the new functions and enhanced features of the proposed IT system as explained in paragraph 10 above, adoption of the computer-assisted multi-modal data collection approach and the use of API for data dissemination under the open data policy of the Government will also help maintain the image of Hong Kong as an Information and Communications Technology advanced and statistically advanced economy, and support the development of a smart city.

Future Use of the 21C IT System

15. To maximise the benefits arising from the acquisition of the IT system for the 21C, the following enhanced sub-systems will be retained for use beyond the 21C operation –

/(a)

- (a) Register of Quarters Sub-system – for continuous updating of a complete list of quarters and building details in Hong Kong which serves as the sampling frame for other household surveys and future population censuses/by-censuses;
- (b) Sampling Sub-system – for sample selection for other household surveys and future population censuses/by-censuses;
- (c) Digital Mapping Sub-system – for maintaining digital maps for supporting other household surveys and population censuses/by-censuses;
- (d) Statistics Dissemination Sub-system – for on-going maintenance of statistical tabulations of 21C results for use in the Thematic Website and IDDS;
- (e) IDDS – for on-going provision of interactive tabulation services on 21C results to the public; and
- (f) 21C Thematic Website – for on-going dissemination of main 21C results to the public.

16. As in the past, we will re-deploy the equipment and facilities of the rest of the sub-systems of the 21C, including the mobile tablets for data collection, to meet other operational needs in C&SD where appropriate after completion of the 21C, with a view to maximising the utilisation of resources available. Any surplus serviceable equipment will then be re-deployed to other government departments or donated to public sector schools and/or non-governmental organisations following the established procedures of the Government. We will also re-use the sub-systems in the 2026 Population By-census as far as possible taking into account the business and operational requirements then.

FINANCIAL IMPLICATIONS

Non-recurrent Expenditure

17. We estimate that implementation of the proposal will incur a non-recurrent cost of \$202.68 million over a four-year period from 2019-20 to 2022-23 for the acquisition of hardware, software and related services. The detailed breakdown is as follows –

/2019-20

	2019-20	2020-21	2021-22	2022-23	Total
	\$'000	\$'000	\$'000	\$'000	\$'000
(a) Hardware	21,573	56,816	-	-	78,389
(b) Software	5,959	13,148	-	-	19,107
(c) Communication network	2,904	3,786	6,754	995	14,439
(d) Implementation services	7,240	16,953	11,879	2,683	38,755
(e) Contract staff	8,844	11,004	2,476	636	22,960
(f) Hosting and network security services	1,143	2,118	2,502	2,116	7,879
(g) Miscellaneous	1,548	1,106	72	-	2,726
(h) Contingency	4,921	10,493	2,368	643	18,425
Total	54,132	115,424	26,051	7,073	202,680

18. On paragraph 17(a) and 17(b) above, the estimates of \$78,389,000 and \$19,107,000 are for the acquisition of hardware and software respectively. The hardware will include mobile tablets, system servers and storage systems, workstations and other peripherals. The software will include operating systems, database management systems, MDM systems and application development tools, etc.

19. On paragraph 17(c) above, the estimate of \$14,439,000 is for the acquisition of network equipment and data communication facilities, rental of data communication lines and acquisition of mobile data services.

20. On paragraph 17(d) above, the estimate of \$38,755,000 is for the acquisition of services for system implementation and enhancement, equipment installation and system configuration, privacy impact assessment services, security risk assessment and auditing services.

21. On paragraph 17(e) above, the estimate of \$22,960,000 is for hiring IT contract staff for system development work (including system design and implementing system enhancement), providing support during production run and monitoring the implementation services by contractors.

22. On paragraph 17(f) above, the estimate of \$7,879,000 is for expenditure for the cloud hosting services and DDoS network protection service.

23. On paragraph 17(g) above, the estimate of \$2,726,000 is to meet miscellaneous expenditure including cabling work, consumables and staff training.

24. On paragraph 17(h) above, the estimate of \$18,425,000 represents a 10% contingency on the items set out in paragraphs 17(a) to (g) above.

Other Non-recurrent Expenditure

25. The proposed implementation of the project will require an IT project team for project management, procurement of hardware, software and services, system analysis and design, site preparation, quality assurance, acceptance testing and implementation support, etc. This will entail a non-recurrent staff cost of \$16,703,000 during 2019-20 to 2022-23. The provision will be reflected in the annual estimates of the respective financial years.

Recurrent Expenditure

26. We estimate that the proposed system will entail an annual recurrent cost of \$0.4 million from 2023-24 onwards, which will be absorbed by C&SD. This expenditure comprises the cost of the retained communication network and hosting service charge.

Cost Avoidance

27. We estimate that the implementation of the proposed system will bring about one-off cost avoidance of \$56,175,000. It covers the savings on (i) the staff and equipment costs for conducting face-to-face interviews with the implementation of telephone interviews; (ii) the cost for manual coding by implementing code-searching functions in the data collection sub-systems; and (iii) the printing, storage and data capturing costs for paper questionnaires with the use of mobile tablets in face-to-face interviews.

Encl. 28. A cost-benefit analysis for the proposed system is at Enclosure.

/IMPLEMENTATION

IMPLEMENTATION PLAN

29. Subject to approval of funding, we plan to adopt the following implementation schedule –

	Major Activities	Target Completion Date
(a)	Phase I (Pilot Survey)	
	- Equipment and service procurement	March 2020
	- System design, development/enhancement and user acceptance test of sub-systems supporting pre-census preparation and census operation	June 2020
	- Pilot Survey	August 2020
(b)	Phase II (Census operation and data dissemination stage 1)	
	- Equipment and services procurement	March 2021
	- Refinement and user acceptance test of Phase I sub-systems	June 2021
	- Census operation	August 2021
	- System design, enhancement and user acceptance test of sub-systems supporting dissemination of summary results	September 2021
	- Dissemination of summary results	February 2022
(c)	Phase III (Data dissemination stage 2)	
	- Implementation of e-service for interactive data dissemination	September 2022
	- Dissemination of other census results	March 2023
(d)	Post-implementation support	March 2023

/PUBLIC

PUBLIC CONSULTATION

30. Extensive consultation on various aspects of the 21C started in July 2018 to ensure that the 21C would be designed and conducted in an efficient and cost-effective manner. Some 200 parties including District Councils, advisory and statutory bodies, selected organisations including non-governmental organisations and trade and industrial organisations, relevant faculties/departments in academic institutions, and government bureaux and departments have been involved.

31. On 7 January 2019, we briefed the Legislative Council Panel on Financial Affairs on our planning work for the proposed 21C and the estimated funding required for IT systems and services. The Panel supported the proposal.

BACKGROUND

32. Since 1961, it has been an established practice in Hong Kong to conduct a population census every ten years and a population by-census in the middle of the intercensal period. The last population census and by-census were conducted in 2011 and 2016 respectively. Following this practice, C&SD will conduct the next population census in 2021.

33. The aim of conducting population censuses/by-censuses is to obtain up-to-date benchmark information on the socio-economic characteristics of the population for studying the direction and trend of population changes. The data are key inputs for making projections concerning population, household and labour force. Such information is vital to the Government for programme planning in areas like education, housing, transport, social services and health services and the formulation of population policies. The information is also important to the private sector for business and research purposes.

34. The difference between population censuses and by-censuses is that the former take a complete headcount of the population while the latter do not. In both population censuses and by-censuses, detailed characteristics of the entire population are inferred from the results of a large sample in accordance with statistical theory. The sizable scale of population censuses/by-censuses, as compared to other general sample household surveys, can provide statistics of high precision even for population sub-groups and small geographical areas.

/35.

35. On 12 July 2014, the Finance Committee approved \$88,814,000 (vide FCR(2014-15)11) for acquiring computer equipment and services for the 16BC.

Financial Services and the Treasury Bureau
February 2019

**Cost and Benefit Analysis for the Acquisition of
Information Technology Equipment and Services for the 2021 Population Census**

		Cash Flow (\$'000)					
		2019-20	2020-21	2021-22	2022-23	2023-24	Total
1	Non-Recurrent						
	Expenditure	54,132	115,424	26,051	7,073	-	202,680
	Staff Cost	2,263	4,951	4,951	4,538	-	16,703
	Total Non-Recurrent Cost	56,395	120,375	31,002	11,611	-	219,383
2	Recurrent						
	Expenditure	-	-	-	-	400	400
	Total Recurrent Cost	-	-	-	-	400	400
	Total Non-Recurrent and Recurrent Cost (A)	56,395	120,375	31,002	11,611	400	219,783
3	Savings						
	Realisable Savings	-	-	-	-	-	-
	Notional Savings	-	-	-	-	-	-
	Cost Avoidance	-	-	56,175	-	-	56,175
	Total Savings (B)	-	-	56,175	-	-	56,175
	Net Savings (C) = (B) - (A)	(56,395)	(120,375)	25,173	(11,611)	(400)	(163,608)
	Net Cumulative Savings	(56,395)	(176,770)	(151,597)	(163,208)	(163,608)	
