
**Implementation of Data Alignment Measures
for the Alignment
of Planning, Lands and Public Works Data**

**Final Report (Volume 2I)
Data Custodianship and License Agreement**

March 2004

Amendment History					
Change Number	Revision Description	Pages Affected on Respective Version	Revision/Version Number	Date	Approval Reference
0	Initial Release		1.0	4 Mar 2004	

Volume 2I – Data Custodianship and License Agreement

Table of Contents

1	Introduction.....	1-1
1.1	Purpose.....	1-1
1.2	Data Custodianship.....	1-2
1.3	License Agreement	1-9

1 Introduction

1.1 Purpose

1.1.1 This document would be the draft version of the Practice Note which advises on the Data Custodianship and license arrangement in data exchange processes in DAM.

1.1.2 This Practice Note shall be read in conjunction with HPLB TC No. xx/xx and ETWB TC No. xx/xx.

Data Custodianship

1.1.3 Data Custodianship plays a major role in spatial data management, in particular in this DAM initiative. It defines and provides for accountability for maintenance of CSUs integrity and identifies the duty of care for Data Agent, Data Owners and Data Users to ensure the integrity, accuracy, validity, quality, timeliness and consistency of CSU data. A well coordinated custodianship will help:

- (a) Eliminating unnecessary duplication in the collection and maintenance of the most common types of spatial information;
- (b) Managing information on behalf of others;
- (c) Providing a sound spatial information infrastructure;
- (d) Assisting the production and management of spatial information products; and
- (e) Facilitating the collection of fundamental datasets and spatial information.

1.1.4 In the context of DAM, PDs will take up different roles in the custodianship. The following principles¹ will apply:

- (a) Execute the trusteeship of the Data Agent – Data Agent does not “own” all of the attributes of the CSUs. They hold the data in trusteeship of the others.
- (b) Maintain standards – The standards would be maintained by the DAM Management Committee. These standards include CSU specification, symbology specification, file formats standard. Relevant

¹ GUIDELINES FOR CUSTODIANSHIP, Permanent Committee on GIS Infrastructure for Asia and the Pacific (PCGIAP), April 2000

policies/Technical Circulars for the implementation of DAM would be maintained at bureau level.

- (c) Maintenance of Information – Enforce the maintenance of information to conform to requirements with respect to its accuracy, validity, timeliness and quality.
- (d) Authoritative source – Define the mandate of the management framework of the Provisional DAS Organisation Structure.
- (e) Accountability – Define the accountability of care and maintenance of data and to ensure its integrity throughout its maintenance period.
- (f) Information Collection - Enforce the collection of information to conform to privacy and sensitive data requirements.

License Agreement

- 1.1.5 In the data exchange process between Data Owners and Data Users through the Data Agent, the Data Agent and Data Users shall conform to the conditions laid down on the license arrangement, when applicable, in particular for issues relating to copyright.
- 1.1.6 Data Agents, Data Owners and Data Users are all required to conform to guidelines for proper data handling in the context of CSU data, e.g. data privacy and data sensitivity, when applicable.

1.2 Data Custodianship

- 1.2.1 The principles of custodianship would help deliver a well-defined accountability with respect to the care and maintenance of CSU data to ensure the data integrity throughout the life cycle in the data exchange process.
- 1.2.2 On implementation of DAM, there would be one Data Agent who is responsible to administer the dissemination and maintenance of each CSU, attributes of which might come from different Data Owners for downstream use by Data Users. The trusteeship and ownership of the data in the exchange process need to be defined.
 - (a) Definition and Roles of the Data Agent, Data Owner and Data User; and
 - (b) Assignment of Data Agent and Owner.
- 1.2.3 Please refer to Volume 2A to Volume 2E for designation of Data Agent, Data Owners and Data Users of each CSU.

Assignment of Data Agent, Data Owner and Data Users

- 1.2.4 In the context of CSU data, the Data Agent, Data Owners and Data Users shall be the designated PD. Unless otherwise instructed by the PDs, the GIS Managers of the respective PDs will represent the department (in the capacity of Data Agent/Data Owner/Data User) to attend all issues relevant to the discharge of the duties. The GIS Manager shall be the one nominated by the PD as per HPLB TC 2/96, ETWB 19/96 and Volume 2H of the Final Report).

Role of Data Agent

- 1.2.5 The Data Agent of a CSU shall take up the responsibility to be the central point of contact for disseminating information about the CSU. The Data Agent, nominated by the respective CSU sub-working group, shall be appointed by the DAM Management Committee.
- 1.2.6 The Data Agent shall be responsible to maintain the required IT infrastructure good for data dissemination purpose. The provided IT infrastructure should conform to guidelines for proper data handling in the context of CSU data, e.g. data privacy and data sensitivity, when applicable. Also, the Data Agent is responsible for data availability to PDs at specified time intervals.
- 1.2.7 The Data Agent is responsible to work with Data Owners to implement the CSU standards. This includes:
- (a) Enforce the specification of CSU - Ensure the data from the Data Owners conform to specification requirements with respect to data completeness, timeliness, symbology standard and file formats standard.
 - (b) Prepare metadata of each CSU and submit to hosting PD of the Metadata Catalogue System.
 - (c) Respond to Data Owners/Data Users requests for enquiries on exchanged data.
 - (d) Issue and maintain CSU IDs - Issue and maintain CSU IDs for the dataset, and ensure the ID's uniqueness to allow PDs to perform translation and matching of their data with respect to the CSU dataset.
 - (e) Administer dataset ownership;
 - (f) Observe license arrangement;
 - (g) Resolve CSU related issues brought up by Data Users and/or Data Owners, if possible or refer the issues to DAM Management Committee if needed.

Role of Data Owner

- 1.2.8 The Data Owner shall be the PD who agrees to take up the responsibility to being a Data Owner of a defined set of attributes, from the common attribute list agreed by CSU sub-working group, to the CSU specification.
- 1.2.9 The data from data owners would come from different sources:
- (a) Data originated from PD – the Data Owner is the exclusive owner of the data prepared for their core business, unless there is separate agreement with other PD who agrees to take up the role of data owner.
 - (b) Data originated from non-PD – the data source government department who, usually under a license agreement, releases data to a PD for a particular business purpose. In the context of CSU and with consent from the data source department, this PD becomes a Data Owner on behalf of the data source department, and would be responsible for preparing the relevant CSU data according to the CSU specification and exchanging the data with other PDs.
 - (c) Data originated from non-government agency – there are situations that a PD is delegated with the authority to receive and register the data from the data source agency (a non-government department). In the context of CSU and conforming to the business rules, this PD becomes a Data Owner who would be responsible for preparing the relevant CSU data according to the CSU specification and exchanging the data with other PDs provided that the delegated authority permits.
- 1.2.10 The Data Owner would be responsible for:
- (a) Preparing the above CSU data according to the CSU specification and ensuring the completeness, accuracy and quality of its datasets and making them timely available for sharing;
 - (b) Cooperating with the Data Agent and other Data Users to resolve issues related to the CSU dataset;
 - (c) Observe license arrangement;
 - (d) Conforming to guidelines for proper data handling in the context of CSU data, e.g. data privacy and data sensitivity, when applicable;
 - (e) Ensuring that the data entity and attributes are described and documented appropriately to the metadata definition; and

- (f) Providing adequate information in consistent structural digital format and support to the Data Agent in the preparation of metadata, which would be performed by the Data Agent.

- 1.2.11 There can be multiple Data Owners for one CSU, who contribute different portion of the CSU data.
- 1.2.12 Despite the fact that the Data Owner is responsible to ensure the completeness and accuracy of its datasets and making them timely available for sharing, the Data Owner could only assume that the data made available to them under the scenario 1.2.9(b) and 1.2.9(c) have been checked correctly by the data source and the update frequency of such data in the CSU dataset will depend on the frequency that the Data Owner can collect the information from the data source. In case there is any doubt about the data, the Data Owner will follow up with the data source through existing normal business procedures and report the result to the Data Agent, Data Users or the provisional DAS organization structure.

Role of Data User

- 1.2.13 Data User would be responsible to facilitate license agreement arrangements for external contractors. The Data User shall undertake or sign a license agreement, when applicable, to ensure that the data obtained from the respective Data Owners through the Data Agent is for the purpose stipulated in the license agreement. Also they shall conform to the requirements with respect to data privacy, security and confidentiality.
- 1.2.14 Data User should report any error, when spotted, to the respective Data Agent who would follow up with the Data Owners accordingly.

Responsibilities and Terms for Re-Disseminating Data

- 1.2.15 In the interim stage, it was recommended in DAM 1 that only simple data storage/retrieval would be considered in the data dissemination system. Details of data dissemination will be covered in Volumes 2A – 2E.
- 1.2.16 In the data exchange process between Data Owners and Data Users through the Data Agent, the Data Agent and Data Users shall conform to the conditions laid down on the license arrangement, when applicable, in particular for issues relating to copyright.
- 1.2.17 Data Agents, Data Owners and Data Users are all required to conform to guidelines for proper data handling in the context of CSU data, e.g. data privacy and data sensitivity, when applicable.
- 1.2.18 The responsibilities of the Data Agent for disseminating data are summarized below:

- (a) Support
 - (i) No help desk service would be provided. Normal enquiry and data dissemination service would be provided during normal working hours by designated staff of the PDs, and
 - (ii) Response to enquiries should be given to PDs within a reasonable time. When prolonged investigation is required, an interim reply is required.
- (b) Cost - PDs should make their own funding application for the designated DAM component under their jurisdiction and their other business initiatives. Service provided by the data agent and data owners will be at no cost to other PDs.
- (c) Term - the current arrangement is that the Data Agent would continue to take on its role unless otherwise instructed by the DAM Management Committee.
- (d) Warranties and Indemnity - the Data Agent should warrant that:-
 - (i) The dissemination system can accurately perform the function with respect to data storage and their conversions;
 - (ii) They will rectify any defects in the data/software once they are made known. If the identified defect is due to data integrity from the data source data, the issue should be referred to Data Owner for follow up. Data Agent will provide necessary assistance upon request;
 - (iii) The dissemination system is adequately provided with effective security measures to maintain the integrity of the dissemination system; and
 - (iv) The dissemination system is maintained regularly.
- (e) When the DAM is extended to include quasi-government and non-government bodies on migration from DAM to DAF, there is a possibility that the Data Agent role could be taken up by a non-government body. In such case, the interest of the PDs should be protected. Government departments (Data Owners/Data Users) should be indemnified when they suffer any loss from the following events (although unlikely to happen when a PD takes up the Data Agent role):-
 - (i) Loss of data and corruption of data arising out of the use of data dissemination system;

- (ii) Where the software used by the Data Agent is in breach of a third party's intellectual property rights causing loss and damage to the Government/PDs;
 - (iii) Where the Data Agent breaches its obligations in relation to the handling of personal data; and
 - (iv) Where the Data Agent breaches its obligation of confidence.
- (f) Resolution of data issues
 - (i) Whenever there is any data issue with respect to its accuracy and validity, conversion issue with respect to its accuracy and specification compliance, and enforcement on the updating mechanism of the CSU data, either initiated by the Data Users or the Data Agents, it should be dealt with direct between the Data Agent and the respective Data Owners. Data Users should be informed accordingly on the resolution of the issues; and
 - (ii) When the issues could not be resolved, it should be referred to the DAM Management Committee through the appropriate sub-working group.

Measures for Ensuring the Currency, Timeliness and Quality

- 1.2.19 Currency is best described as how up-to-date or current the data is when compared against the data's captured date, or the date it is to be used. Data quality in terms of the currency erodes over time. Data that may have been an accurate representation at a certain point in time could become invalid if measured against latest current time.
- 1.2.20 The requirements for currency, timeliness and quality is defined in Volume 2A to Volume 2E. The quality description would also be described in the respective metadata of the CSUs. Such requirements would take into the genuine business need, cost impact to the Data Owner for the marginal improvement to its currency, timeliness and quality.
- 1.2.21 On notification of any issues from the Data Users, Data Agent would deal direct with Data Owners at sub-working group level. If there is any issue that cannot be resolved, it would be referred to the DAM Management Committee.
- 1.2.22 Data currencies, timeliness and quality are major requirements from Data Users, but the arrangements need to be agreed with respective Data Agents and Data Owners on the future data exchange process of the CSU, which is covered in detail in Volumes 2A to 2E. To ensure that the requirements can be met, a number of measures are recommended for the different stages involved in:-

- (a) Data collection and submission;
 - (i) A future workflow for each CSU is available to describe the updating processes for the CSU throughout its lifecycle in DAM1.
 - (ii) The Data Agent, together with the Data Owners and Data Users, would agree on an updating frequency schedule which describes the updating frequency and time interval for which a data subset (in textual attribute or spatial form) would be made available. In deriving the updating and maintenance plan, acceptable limits on the quality of their information with respect to accuracy, timeliness, time from capture to usage, and etc are defined and established.
 - (iii) The Data Owner is required to establish internal procedures and guidelines, where applicable, to prepare data according to the CSU specification and provide regular and timely update of data to the Data Agent via the agreed channel(s) of submission.
 - (iv) The Data Agent is recommended to establish a data management system, which lays down the procedures for the collection, administration and handling of data received from Data Owners. To ensure the quality requirements are met, a number of measures are recommended as follows:-
 - (1) Logging and monitoring of data submitted by Data Owners;
 - (2) Verification on the validity and compliance of data according to the CSU specification;
 - (3) Investigation and follow up with Data Owners on suspected erroneous data;
 - (4) Corrective actions to be made by the Data Agent/ Data Owner as required.
- (b) Data integration and preparation
 - (i) Data integration and preparation involves the processing and conversion of data acquired from Data Owners to conform to the requirements specified in the CSU specifications. Transformation of data from the data source formats to the spatial file formats requested by the Data users may also be required.
 - (ii) The Data Agent is recommended to establish internal procedures and guidelines for data preparation and compliance assurance of datasets, ensuring the uniqueness of the CSU ID and conformance

to the CSU specification. To ensure the compliance requirements are met, a number of measures are recommended as follows:-

- (1) Perform compliance assurance and integrity checks on the processed and converted geospatial dataset;
- (2) Provide time-stamping or status updating on the CSU;
- (3) Generate metadata for the CSU on a new major release.

(c) Data dissemination

- (i) The distribution of CSU datasets can be made via different channels, such as through postal mail (CD-Rom, tape, and hardcopy) or electronic means such as File Transfer Protocol (FTP), Email, Data Dissemination System for on-line or off-line retrieval, depending on the arrangement to be agreed upon.
- (ii) The Data Agent is recommended to establish procedures and guidelines for data dissemination. To ensure the timeliness and quality requirements are met, a number of measures are recommended as follows:-
 - (1) Responding to requests or making data available to Data User(s) via the agreed channels within an acceptable timeframe;
 - (2) Publishing the latest metadata of the CSU.
- (iii) Whenever the definition of the CSUs is revised, the metadata should be revised or re-generated according to the revised definition. The metadata requirements should cover the content, quality, condition, and other characteristics of data involved in the exchange process. A clear definition will help stakeholders understand what is being exchanged and the content as is relevant to them. In addition, this is a prelude to the consolidation and capturing of metadata into a common Catalogue.

1.3 License Agreement

Measures for privacy protection and sensitive data

- 1.3.1 Data privacy law in Hong Kong is governed by the Personal Data (Privacy) Ordinance, Cap. 486 ("PDPO"). It regulates the collection, storage and use of data related to living individuals from which it is reasonably practicable to identify the individuals. Such data are those that describe an individual and attribute things to an individual so that others can identify a particular individual. It applies to

Data Users in Hong Kong, whether they are individuals, private companies or public bodies.

1.3.2 A “Data User” is defined in the PDPO as a person who, either alone or jointly or in common with other persons, controls the collection, holding, processing or use of the data.

1.3.3 Personal data means any data:-

- (a) Relating directly or indirectly to a living individual;
- (b) From which it is practicable for the identity of the individual to be directly or indirectly ascertained; and
- (c) In a form in which access to or processing of the data are practicable.

1.3.4 The Ordinance aims to control the collection, holding, processing and use of Personal Data. In the context of exchange of CSU data, the Person Data includes the relevant information collected and recorded, whether on computer or manually, in relation to the attributes defined in the CSUs.

1.3.5 In the context of five CSUs discussed in DAM 1, all common data attributes contained in the CSU specification are neither privacy data nor sensitive data, with the status outlined in Table 1.

CSU	Status on data privacy issues and data sensitivity issues
Building	There is neither privacy data nor sensitive data with respect to the attributes of the CSU definition.
Slope	There is neither privacy data nor sensitive data with respect to the attributes of the CSU definition.
Lot	The attribute “Purchase Name” is a personal data. It is confirmed that “Purchase name” would not be included as an attribute.
TPU/SB	PlanD advised that the projected population figures after 7 th + year would be classified as restricted data. It is confirmed that the data after 7 th + year would not be included in the CSU.
Road Centreline	There is neither privacy data nor sensitive data with respect to the attributes of the CSU definition.

Table 1 Status on data issues

1.3.6 As a general requirement, PDs will have an obligation to keep the data for exclusive use unless otherwise permitted by the license condition.

1.3.7 PDPO is not applicable in the current context of the five CSUs discussed in DAM 1, however, it is to note the following for compliance, when applicable on implementation of new CSUs in the future:

- (a) In the situation when multi-ownership attributes are stored by the Data Agent, there is no law in Hong Kong relating to computer security though the law does require that security measures be in place for handling personal data pursuant to the PDPO. The Data Agent cannot make unauthorised use of personal data and the data dissemination systems must be adequately secure to ensure safe transmission of personal data. In this respect, measures such as passwords, firewalls and encryption, when appropriate, are common protection mechanisms used. Apart from personal data, confidential information should also be treated with care.
- (b) Confidential information can be broadly classified into three categories: personal information, government secrets and trade secrets. If information is disclosed for a limited purpose, then the recipient of such information (i.e. Data Users) will be under a duty to keep it confidential. Obligations of confidence may also be imposed through the license arrangement. The obligation of keeping electronically delivered information confidential is likely to be discharged if (1) the computer systems are adequately secure and (2) all steps have been taken to ensure that there is no unauthorised use or disclosure of confidential information.
- (c) If any of the services is maintained and/or provided by sub-contractors, all such approved sub-contractors must sign a confidentiality agreement in substantially the same terms as imposed on the Data Agent in the contract.

1.3.8 Some PDs would need to handle personal data for their business. These personal data are normally exclusively used within the department, e.g. name, and address of household's data by RVD and LR; and ownership information in LandsD and BD. PDPO would apply to individual PDs in the handling of personal data.

1.3.9 The PDPO is founded on six data protection principles (DPPs), which state the requirements on how personal data should be handled unless the Ordinance allows an exemption. Individuals are not exempt from compliance with the DPPs except where they collect and use personal data only for the management of personal, family or household affairs or for recreational purposes.

Standard license conditions

1.3.10 It has been a practice that Data Users would need to sign a kind of license agreement prepared by Data Owner (a government department). The signed agreement will mean confirmation of the acceptance of the terms and conditions laid down on the agreement by the Data Users. For example, license agreement prepared by LandsD for the use of B1000 data. There are different license arrangements for Data Users from non-government organizations.

- 1.3.11 In the context of CSU, there are some CSU attributes whose source is from a non PD. E.g. building information of HA properties for the Building CSU. Although PlanD offers to take up the middleman role to distribute the data to other PDs, consent from data source, if not a PD, must be sought and acknowledged in the license agreement between HD and PlanD in this case.
- 1.3.12 The license agreement would come in different forms:
- (a) PlanD
 - (i) PlanD would simply include a clause in their covering letter or email: "Please be reminded that the above data are for your internal use only and they should not be reproduced or distributed in part or in whole and in whatever forms and media to a third party without the consent of this Department."
 - (b) LandsD
 - (i) If in case the digital data is distributed to a department's consultant/contractor, the department and its consultant/contractor have to complete undertaking form (SMF-0095 & SMF-0096) and agree to be abided by its terms and conditions.
 - (ii) If the data is for internal use, the department does not need to complete a form, but need to declare in the requesting memo, that data is for internal use within the dept. LandsD does not know whether there are guidelines/code of practice that the staff of these departments have to follow. It is up to the section head who ensures that data is handled properly.
 - (iii) A set of 1:1000 BMS data together with full set of 1:5000, 1:10000 & 1:20000 digital topographical maps in DGN format and the DOP5000 digital orthophoto data have been provided to CED, HyD and DSD. Specific block consent has been given to CED, HyD and DSD for distribution of these data to their consultant/contractor under specified terms and conditions. Separate consents should be obtained from LandsD for further distribution of the data to other external parties.
- 1.3.13 It is recommended that similar license agreement SMF-0095 & SMF-0096 should be used for DAM purpose.