CITY OF SYDNEY

REQUEST FOR TENDER - RFT 931/09
DIGITISATION OF ARCHIVAL RECORDS 2009-2012

SPECIFICATION

Item	Description of service to be supplied
1.0	**DIGITISATION OF GENERAL CORRESPONDENCE & SUBJECT FILES (approx 2000 onwards)** In this section 'processing' means the creation of digital copies of folders (= 'files') of documents. A file will contain multiple documents ranging in size from smaller than A5 up to occasional A0. Most documents are A4 size. Folders will typically contain a mix of single and double sided documents. Most documents are text, but contents may also include, for example: Photographic formats, CDROM/DVD disks, maps and plans. Plans in folders may be 'attached plans' meaning bound into the folder; or 'separated plans' meaning loose inside the folder (normally folded inside the folder cover). <u>Attached</u> plans are to be digitised alongside the other documents in the folder. <u>Separated</u> plans are to be digitised as a separate PDF file. Folders will require minimal preparation and reassembly. After scanning the paper documents can be replaced loose in the folders, without rebinding, and the folders replaced <u>spine down </u>into the archives boxes in which they have been stored. Digital files are to be multi page PDF, one file per folder unless otherwise required. Resolution is to be 300dpi for all original document formats. OCR required enabling *.PDF content/string searching with Adobe Reader. The order/sequence of the scanned pages is to be the same order/sequence as in the original folder, including colour, greyscale, and 'attached plans'. Scan in black and white <u>except for</u> – (1) Documents with meaningful colour are to be scanned in colour.

	Example: drawings with a colour coded key or legend. Coloured letterheads and coloured date stamps are not meaningful. (2) Documents that are grey shaded should be scanned greyscale. Documents that are: microfiche, 35mm slides, negative film strips, or disks should be checked with the nominated contact officer at City of Sydney to confirm if they are to be scanned. If so, to be scanned as colour using JPEG image compression and exported as .PDF.
1.1	Folder and document preparation, and also reassembly and replacement into archive boxes after digitisation. *Price per hour …*
1.2	Scanning of folders of documents per specification above *Price per thousand pages …*
1.3	Scanning of documents requiring special treatment (for example: microfiche, 35mm slides, negative film strips, or disks) *Provide details of pricing where your price at 1.2 above would not apply …*
1.4	Add metadata to each PDF file comprising elements derived from data supplied by the City of Sydney. Source data will be supplied in electronic form as an XLS spreadsheet – or alternative instructions will be given to enable the metadata to be derived from information shown on each folder. Create a metadata file (as XLS) containing data for all digital image files returned on a disk. Metadata for each file may comprise a total of up to 512 characters, but will usually be no more than 256. *Price per folder …*
2.0	**DIGITISATION OF APPLICATIONS PLANS** In this section 'processing' means the creation of digital copies of original paper plans ranging in size between A3 and A0. Almost all plan sheets are A1 or A0. Some may be equivalent pre-metric sizes. The original plans are rolled and stored in boxes. A box of plans contains on average 30 plan sheets however this may vary. All processing can be assumed to be B&W originals only unless

	otherwise specified.
2.1	Preparation of rolled plans for processing, by flattening, removal of staples and pins etc. Includes assessing sheets to see if any colour scanning is required. *Price per hour ...*
2.2	Digitally scan original paper plans up to size A0 at resolution of 200dpi or better so as to allow reproduction (printing, screen resolution) back to original drawing size without loss of quality. Each set of plans (normally = contents of one box) to be captured as a single multi-page file. File format to be PDF. *Price per plan sheet ...*
2.3	Scan in colour, either whole sets (rolls) of plans or selected sheets within a roll. Scanning in colour is necessary in all cases where colour has been used in the original drawing, or added later as an annotation, and viewing in colour is necessary to interpret the plan. *Price per plan sheet ...* *Please include explanatory notes if there is a price difference between scanning in colour an entire roll of plans, and scanning only selected sheets within an otherwise B&W roll.*
2.4	Add metadata to each PDF file comprising elements derived from data supplied by the City of Sydney. Source data will be supplied in electronic form as an XLS spreadsheet. Create a metadata file (as XLS) containing data for all digital image files returned on a disk. Metadata for each file may comprise a total of up to 512 characters, but will usually be no more than 256. It will be able to be 'cut and paste' from source data supplied by the City. *Price per file ...*
3.0	**DIGITISATION OF ASSESSMENT BOOKS** In this section 'processing' means the creation of digital copies of original paper volumes.

Each image will normally be of a double-page spreads of a volume. Bindings are not tight and no problems are anticipated in creating flat openings for the scanning process. Some volumes up to 60mm thick but most are less than 40mm.

Typical dimensions of an opening to be scanned are: H 470mm x W 580mm. Number of openings per volume ranges from approx 30 to approx 200.

The openings in each volume are sequentially numbered starting with '1'.

There will often be pages with data (eg an index) either preceding or following the run of numbered openings.

It is essential that all page openings with data are scanned and that the file tagging specification below is strictly adhered to.

Each scan must be created as a separate image file. So the product of scanning a given volume will be a set of images each recording a single page opening of the book.

Image files must be tagged according to the following rules:

1. Any preliminary page openings (occurring before the main sequence of numbered pages) should be tagged as CSAnnnnnn_Prelim1 and so on. Where nnnnnn is the barcode ID number found as a bookmark on every volume. Note that the numerals following the CSA prefix are always six digits.

2. Similarly any scanned pages that follow the numbered page sequence will be tagged as CSAnnnnnn_Post1 and so on.

3. Each image created from the main sequence of numbered pages in a volume will be tagged CSAnnnnnn_ppp where ppp is the page number, expanded where necessary to three digits.

It follows from this that an essential task, as part of the scanning process, is to confirm where the main numbered page sequence commences and ends.

The pages are numbered in the top right corner of each double-page opening. Sometimes the number will be contemporary (done at the time the book was made), sometimes it will have been added later, usually in [square brackets] in pencil. Sometimes each page (left and right) will be numbered – in such cases use the odd numbers.

All digital images created from a volume must be placed in a folder, with the CSAnnnnnn number of the volume as the folder title.

	Blank pages containing no data should not be scanned (even if they are numbered).

All pages of a volume should be checked as sometimes a sequence of blank pages will be followed by a page containing data.

Image format must be PDF.

Scanning must be in colour.

Resolution must be 240dpi.

Digitisation must be direct from the volumes, not from intermediate microfilm.

You are given the options (below) of quoting per image or time-based. |
| 3.1 | Digitisation of Assessment Books per specification above:

Price per digital image ... |
| 3.2 | Digitisation of Assessment Books per specification above:

Price per hour ... |
| **4.0** | **DIGITISATION OF LETTERS 1843-1899**

In this section 'processing' means the creation of digital copies of original paper documents. Documents include handwritten and printed papers, and plans and drawings.

The documents are enclosed in manilla folders within standard 'Type 1.1' archives boxes. Each box contains approx 130 folders.

All papers within a folder are considered to be part of the same document regardless of size and contents.

A document /folder will typically contain between 1 and 3 pages but could be as large as 30 pages.

Document size ranges from A6 to A3 folded, equivalent pre-metric sizes. Most are text (handwritten) but there are also plans.

Some documents (especially plans on drafting tissue) are exceptionally fragile. All documents must be treated as fragile and handled with care.

Any document may require preparatory work , especially if it has become dog-eared or small pieces have folded over along the edge, |

hiding some of the writing. Occasionally a document will have to be unfolded.

Handwriting is often very small and somewhat faint, and includes both ink and pencil.

When scanning a document all pages containing any writing or other meaningful data must be digitised. Blank pages or pages with only meaningless marks (eg ink blots) should not be scanned.

Small documents may be scanned as double page spreads.

Each document (= total contents of a manilla folder) must be scanned as a single multi-page image file.

Image files must be tagged according to the following rules:

1. Each manilla folder is stamped on the cover with a unique three-part number that identifies the document. The first component of the number is always '26'. The digital file must be tagged with this three-part number. Note that –

   - The three parts of the number must be separated with underscore symbols.
   - The digits comprising the number must be exactly reproduced in the file tag.
   - The final symbol in the number will sometimes be a special character (such as 'a' or '½'. These must be reproduced in the image file tag.

   Examples of possible image file tags:
   26_1_001
   26_033_1047a

2. The high-res digital images should be tagged with the above numbers. The low res images should have identical tags with the addition of an 'L' prefix. So for example: L26_1_001; L26_033_1047a.

3. All documents digitised from the same archives box must be placed in a folder, and the folder must be tagged with the CSAnnnnnn box number. Where nnnnnn is the barcode ID number found on the box. Note it is always 6 digits after the CSA prefix.

4. Create one folder for the High res images, and one for the Low res images, for each box. So there will be two folders, each with the same box number tag. These should in turn be placed within folders labelled "High res" and "Low res"

Image format must be PDF.

	Scanning must be in colour. Resolution of digitisation must be 240dpi. The 240dpi images (high resolution) must be auto copied to 150dpi (low resolution) and two images accordingly supplied of each document.
4.1	Digitisation of Letters per specification above. Includes all necessary preparatory work. *Price per digital image ...*
5.0	**DIGITISATION OF PHOTOGRAPHS** The City currently utilises specialised software 'Perfect Pictures', vended by Search Tech Pty Ltd, to store and provide access to digital copies of photographic images. The Perfect Pictures software is known as ArchivePix in the City of Sydney environment. The public interface can be seen at http://www.cityofsydney.nsw.gov.au/AboutSydney/HistoryAndArchives/Archives/ServicesForResearchers/SearchTools.asp We have an ongoing requirement to prepare original photographic images for loading into ArchivePix by – (1) Creating low resolution thumbnail, medium and large size images for public web access. (2) Where necessary also creating a high resolution digital image for nearline storage and on-demand access. (3) Adding all necessary metadata tags (4) Uploading images into ArchivePix. Original images format includes: paper prints, negatives, transparencies in various sizes, and digital images. May be colour or B&W. Original material will normally be uplifted from Town Hall House, 456 Kent St, Sydney 2000. Completed work will normally be returned there, and uploading tasks performed if not achievable remotely.
5.1	Performing all the above tasks in cases where creation of a high

	resolution digital copy is NOT required *Average price per image …*
5.2	Performing all the above tasks in cases where creation of a high resolution digital copy IS required *Average price per image …*
6.0	**GENERAL** This section refers to Specification sections 1 to 4 only, above.
6.1	Uplift boxes of records from the Government Records Repository (GRR) at Kingswood and transport to worksite for digitisation processing. Return boxes to GRR when completed. Includes all necessary liaison with GRR to arrange convenient collection/delivery times. *Price per round trip, for a load of minimum 50 boxes and/or volumes …*
6.2	Burn digital images and associated metadata to disk (DVD) and deliver to City of Sydney Archives. *Price per disk …*