Falmouth History Archive Database

ArchiveDb version 2.50

Contents

Falmouth History Archive Database Part 1

ArchiveDb version 2.50

What's new

- Export datagrid
- ISAD(G) compliant fields added to folders and files
- New shortcut icon
- Client-side codebase upgraded to current long term support version 8.0
- Server-side Firebird RDMS upgraded to current version 5
- Performance improvements
- Minor user interface changes

Archive Structure

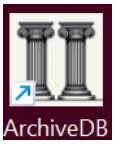
'**Fonds'** or 'Archive Group' is a term often used to describe the whole collection held by an organisation. This may then divide into sub-fonds. Below this, data is collected into '**Series'**. Series may have sub-series. Series hold groupings of similar data. Below series come '**Files'**. A file contains the individual records that relate to a single topic.

ArchiveDb holds data in hierarchical form, which is a tree-like structure, to allow it to represent the Series held by the archive. In the screen print below, '*Falmouth History Archive*' represents the Fonds and the folders beneath it in the tree represent the Series.

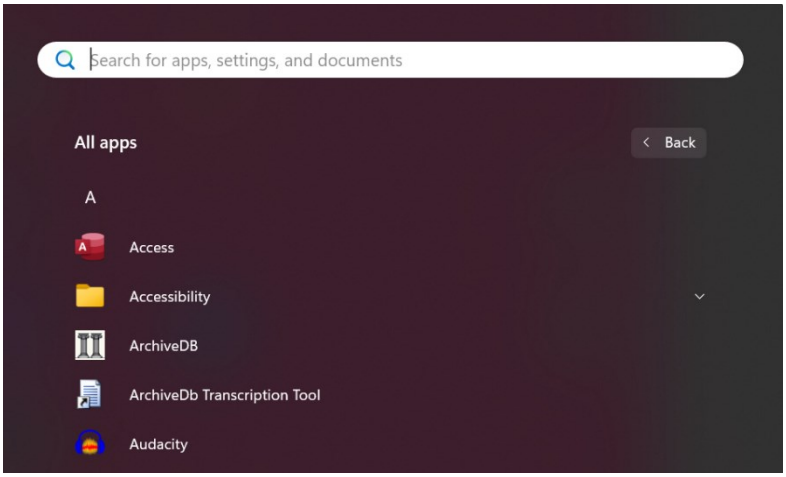
ArchiveDb

Opening the program

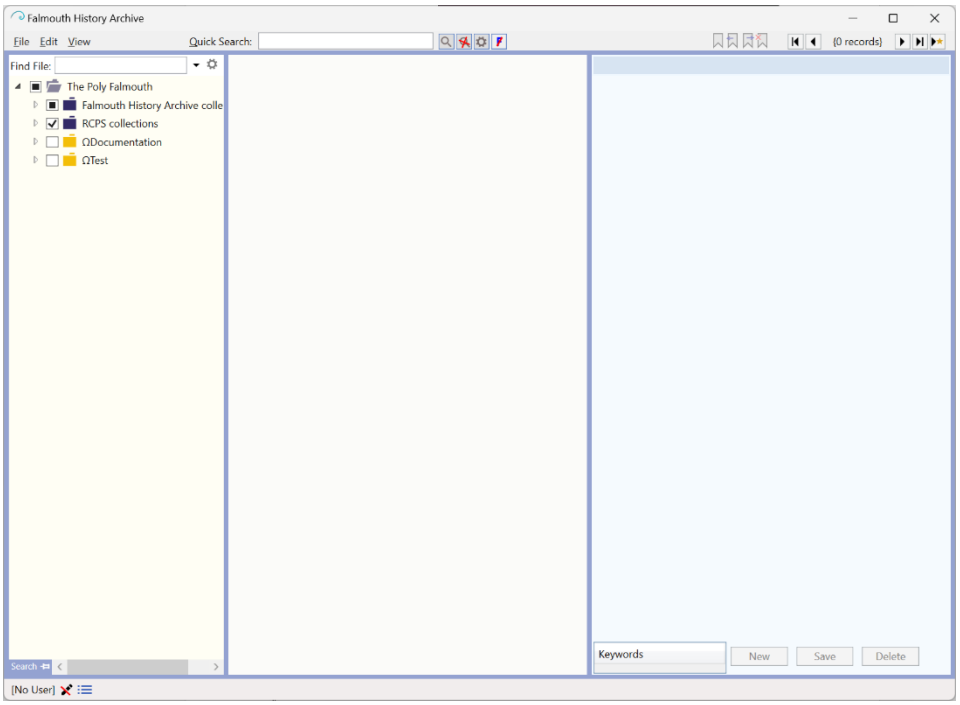
The program can be started by double clicking its desktop shortcut:



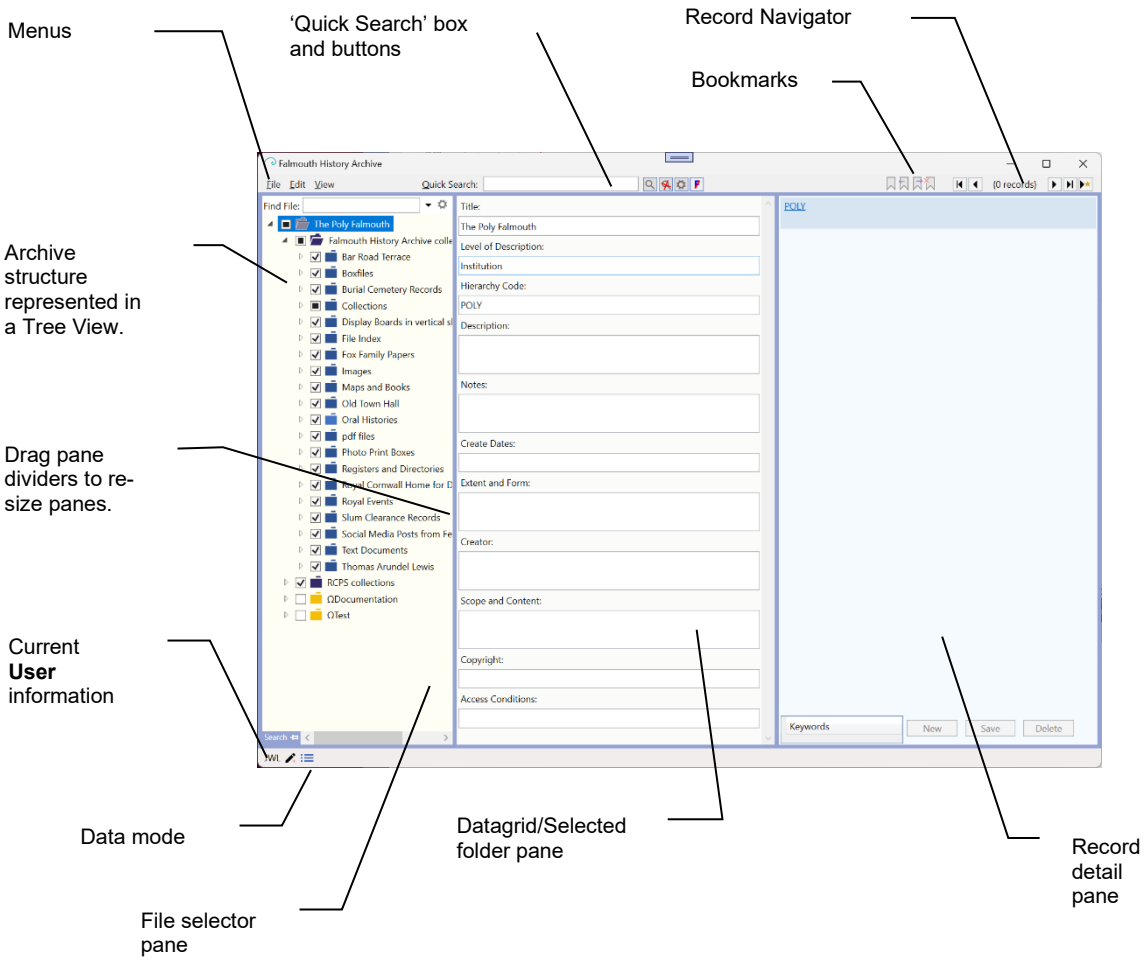
Or by single clicking ArchiveDb on the Start Menu.



This is the main window at start-up

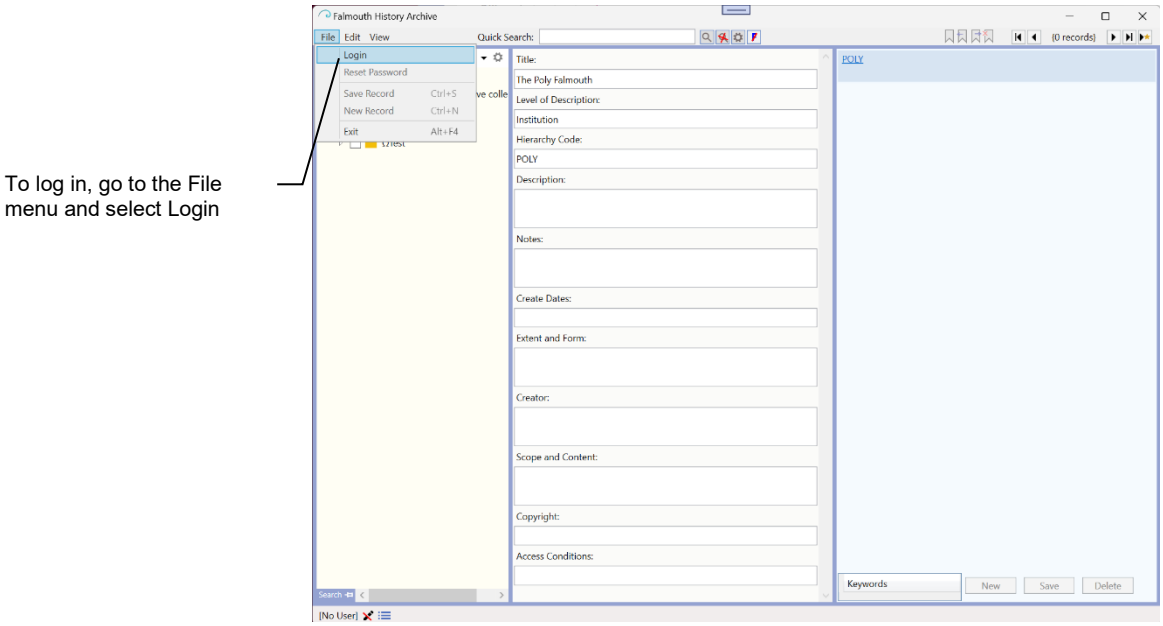


Main window overview

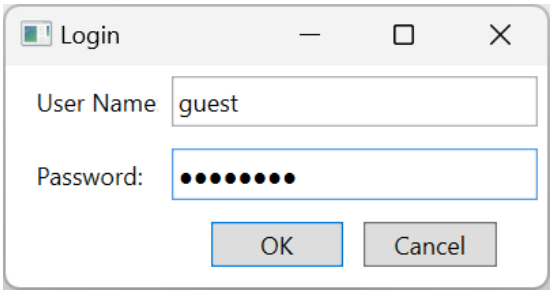


Log in and Privilege

It is not necessary to log in to view data in read-only mode.



Type your **User Name** and **Password** and click OK. Password is case sensitive, User Name is not.



User access depends on user privilege level.

There are 3 privilege levels:

- Read Only
- Editor
- Administrator

Read only access does not allow any data to be added, deleted or modified. Images can be viewed but not saved or downloaded and image properties are not available to prevent a user easily finding images to copy via the file system.

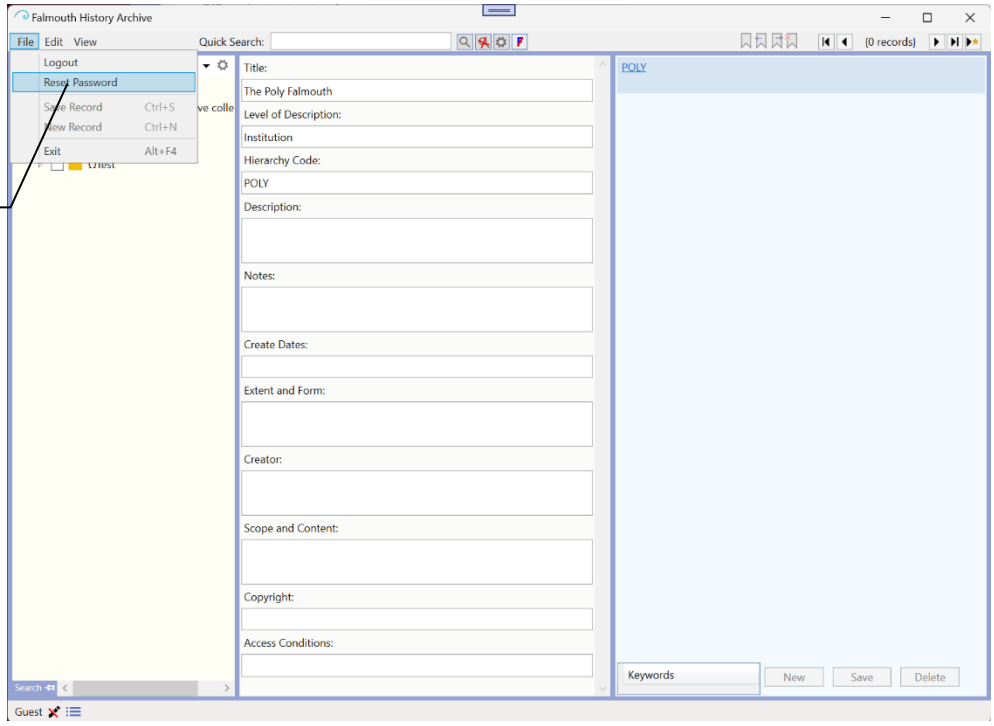
Editor privilege allows the user to add, edit and delete individual records, add and edit Files and also to download image files.

Administrator privilege additionally allows Files to be deleted and Series to be added, edited and deleted, along with other database maintenance features.

By default, all Files have all-user access. This can be modified in File set-up to limit user access to certain Files. See later.

Password Reset

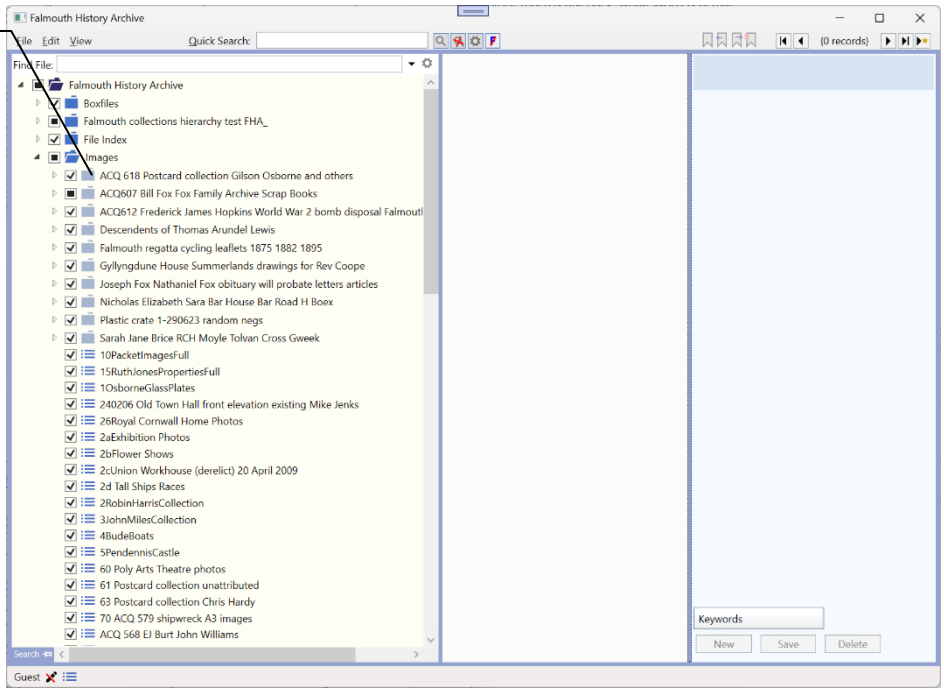
You can change your password by using the ***Reset Password*** command on the **File** menu.



Read Only Mode

Colour Coded Hierarchy

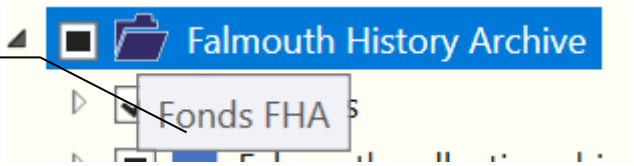
Folder icons are colour coded by their **Hierarchy Level.**
Hierarchy Level corresponds with ISAD(G) ***Level of description***.



Institution: Fonds: Sub-Fonds: Series: Sub-Series: Unspecified:

File:

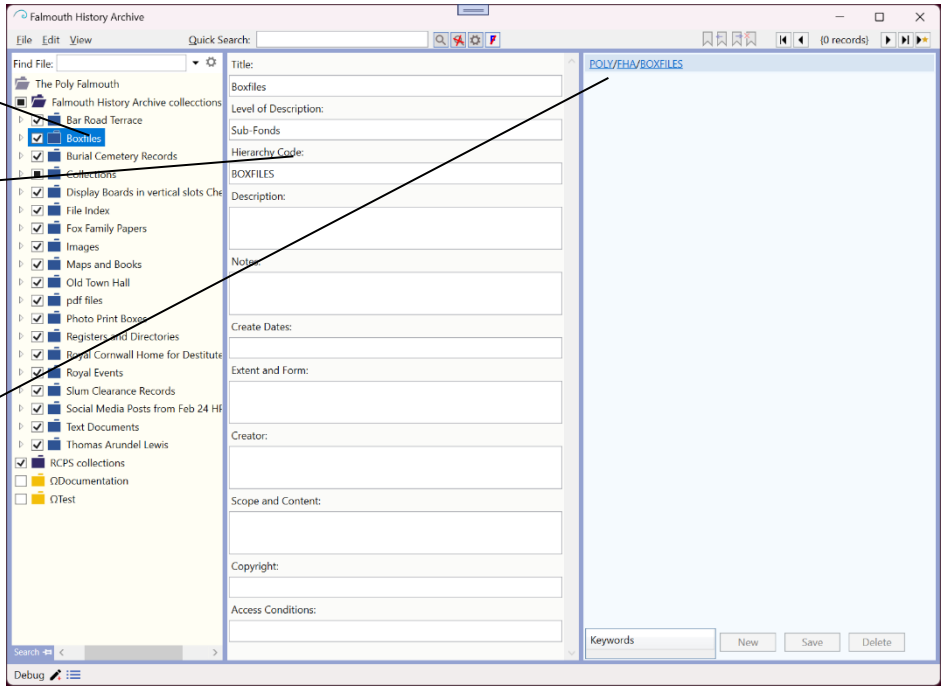
Hover the mouse over the folder icon to display the tooltip showing **Hierarchy Level** and **Code**



Select a Folder…

…to display folder information in the middle pane

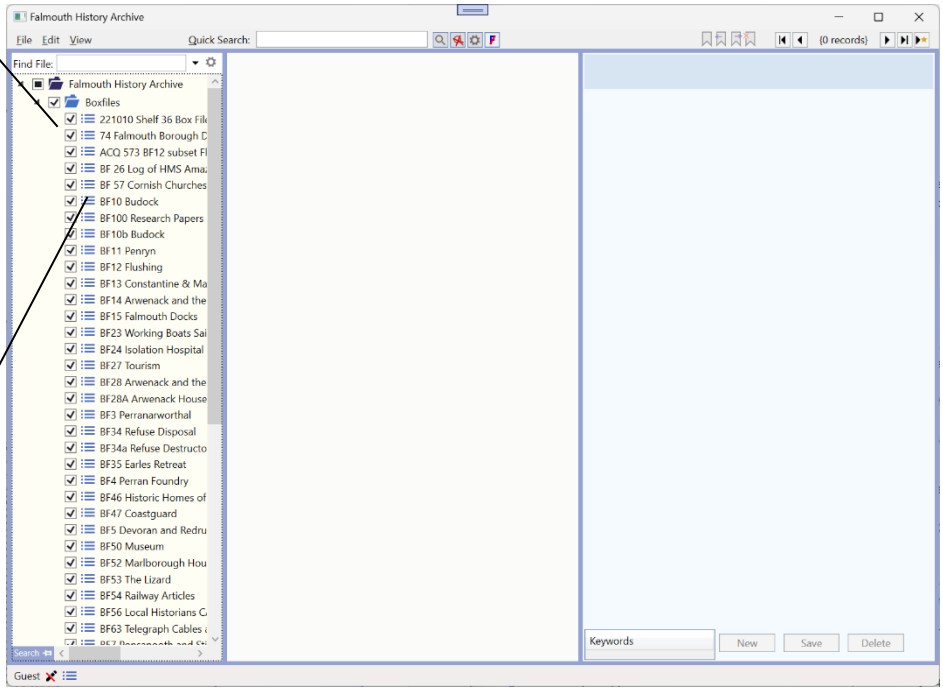
The **hierarchy reference** of the selected object is displayed at the top of the details pane.



Viewing data: Browsing

The Tree View displays **Series** as folders, just like folders in a computer File Explorer. These can be expanded by clicking the glyph or double clicking the folder.

These are the individual **Files** within the *Boxfiles* **Series**.
To browse the content of a file, single click the file name.

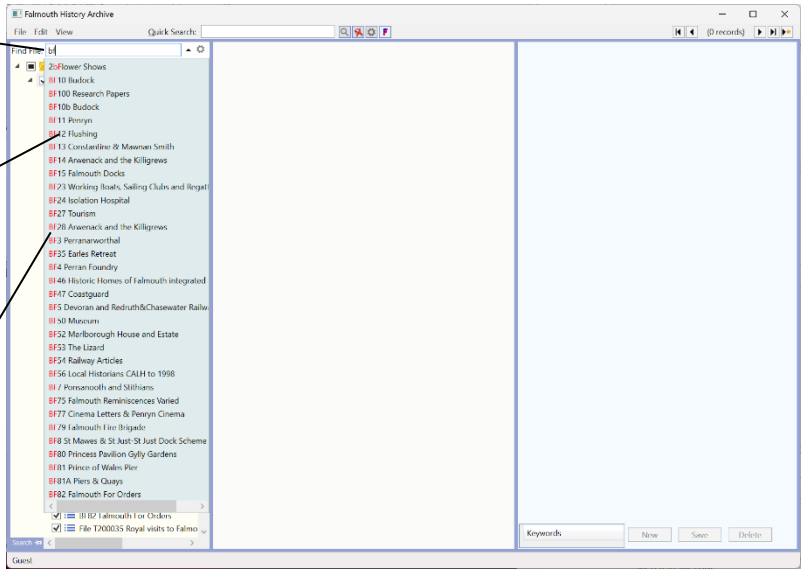


Find File by file name or File keyword

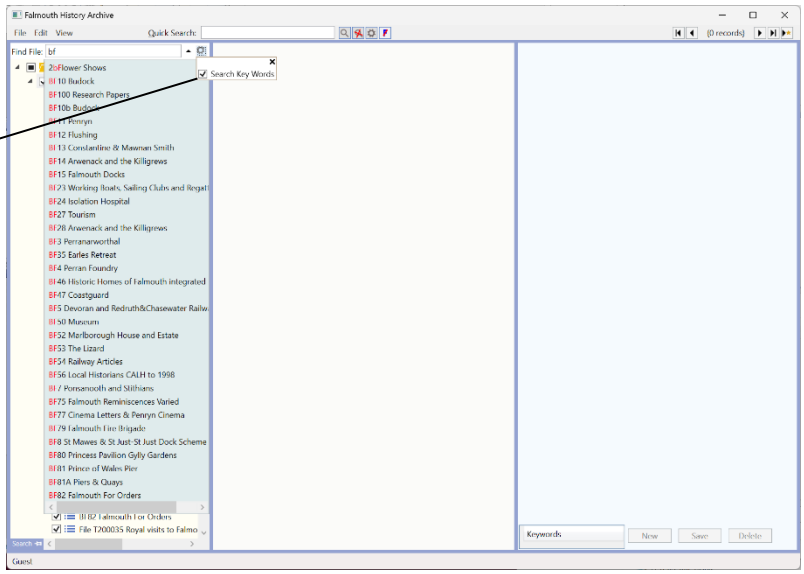
Type some text into the Find File criterion box…

…this will display a list of Files whose file name contains the entered text. The occurrence of the text criterion is highlighted in red.

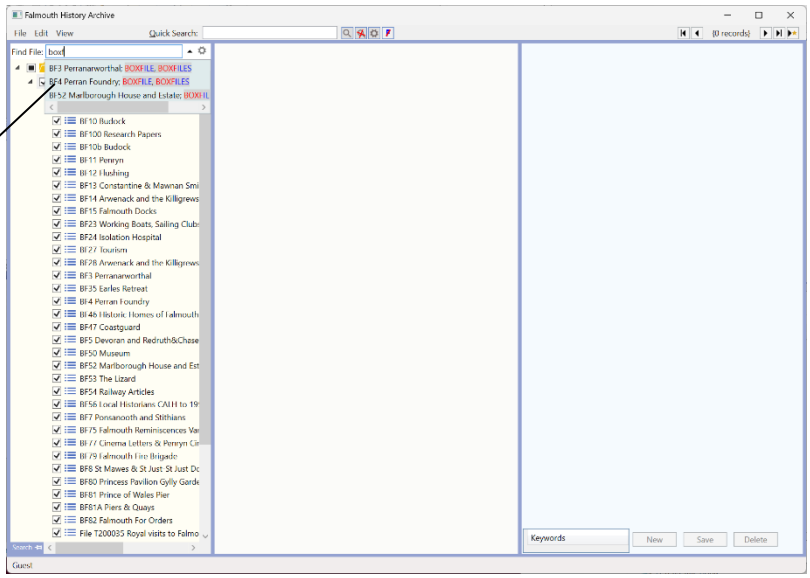
Click a File in the list to Browse its contents



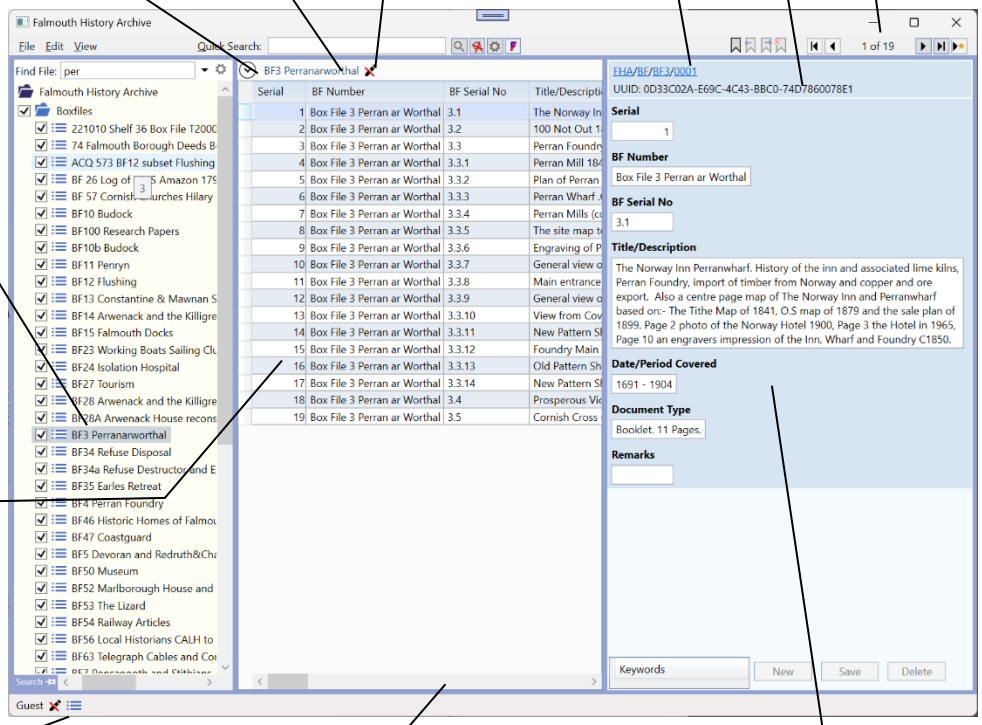
You can opt to search/not search the File keywords. Click the options cog to display the pop-up and check/un-check the **Search Key Words** checkbox



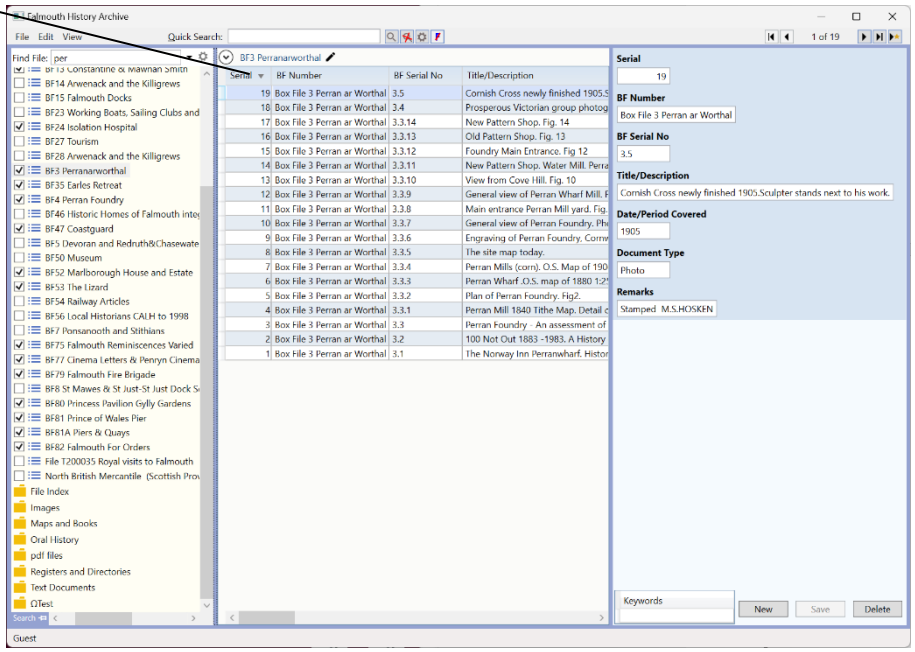
Keyword matches follow a semicolon at the end of the file name. Keyword matches have a grey background. Matched letters are red and unmatched letters are blue. Click a match to Browse the file.



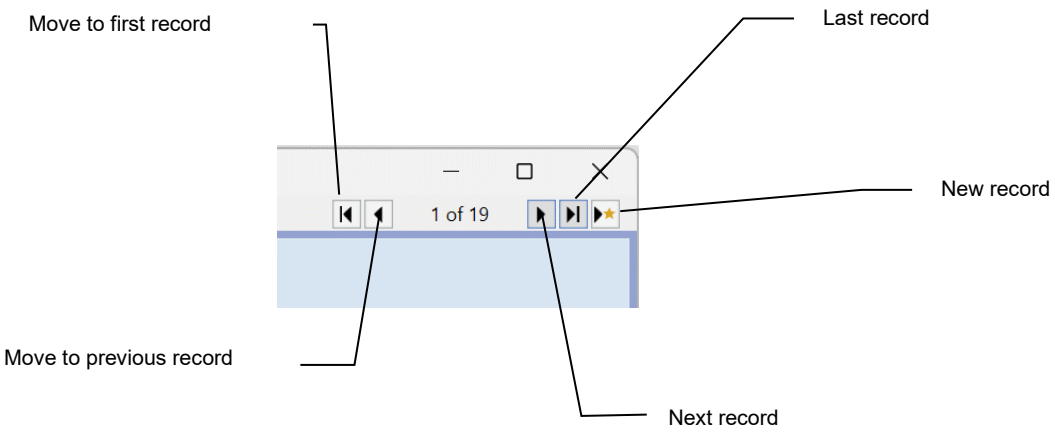
Browse View in Read Only mode:



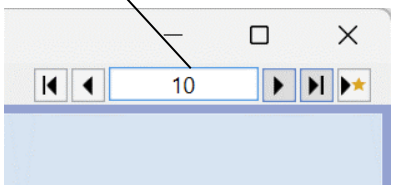
Records can be sorted by clicking the column headers. In this case the column is sorted in descending order. **Long Text** and **Location** columns cannot be sorted.



Record Navigation



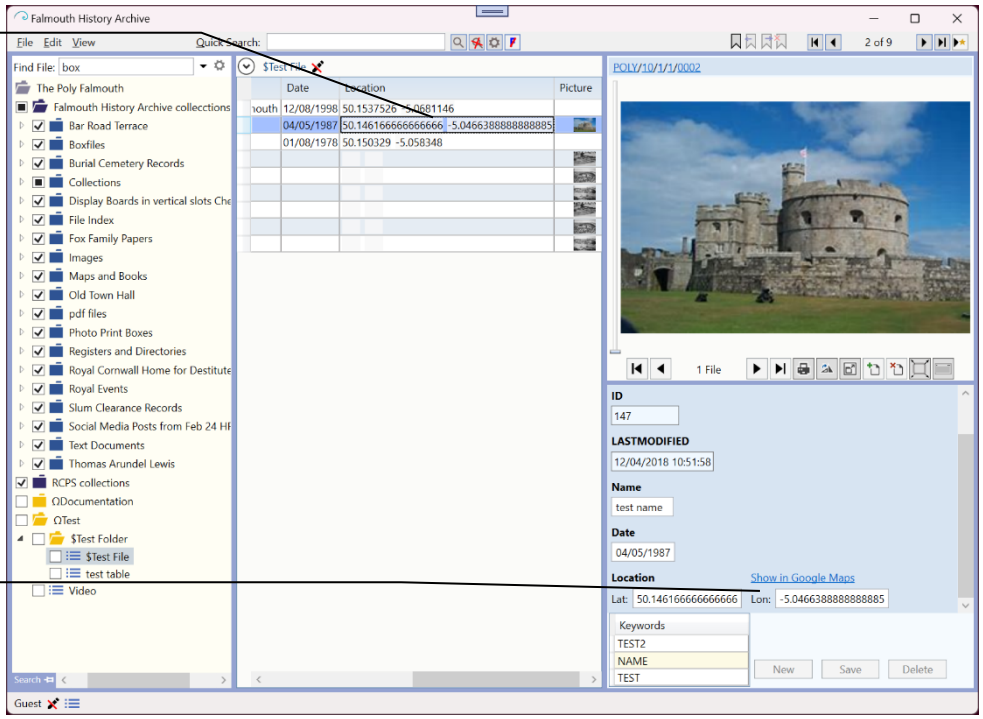
Click into the record number box, type a number and press return {Enter} on the keyboard to go to the record number



Note: these record numbers only represent the ordinal record number within the data grid at the time it is loaded. They do not represent a searchable entity and are subject to change as records are added or deleted.

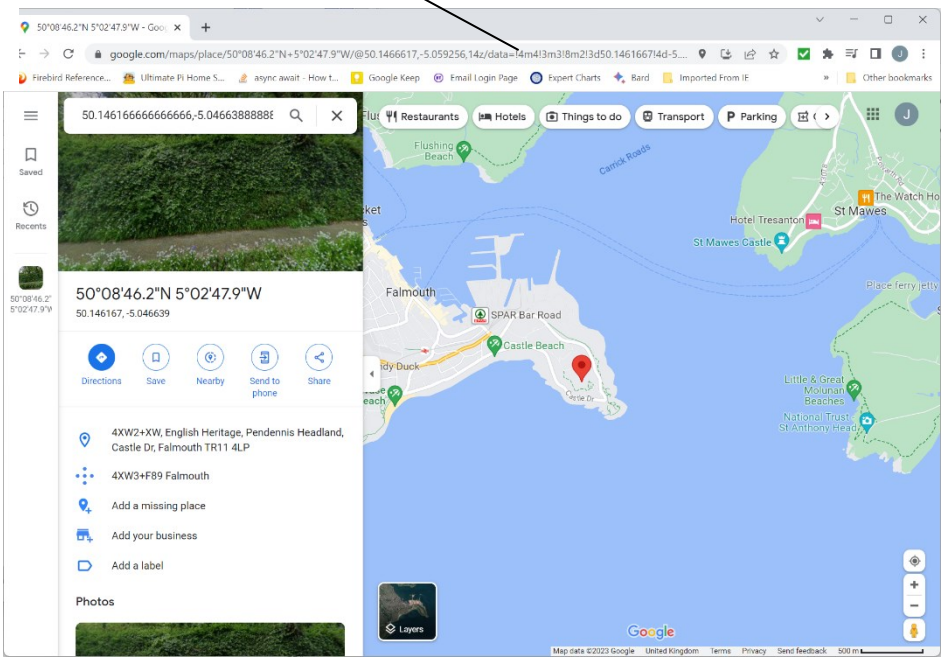
Files with Geo-location data

Geo-location data is displayed in decimal fraction format in the grid.
Northerly latitudes are positive, southerly are negative, easterly longitude is positive and westerly negative.



The record viewer displays the composite Lat/Lon field as shown. Additionally, there is a hyperlink to display this location in the computer's default web browser using Google Maps

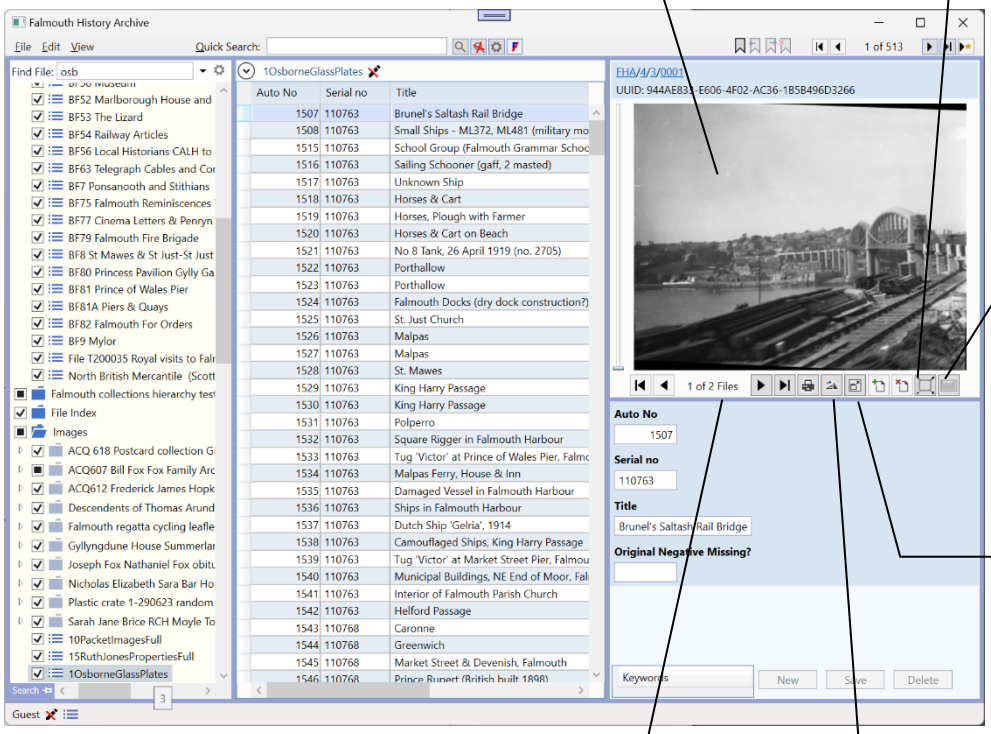
This is how the location will be displayed in Google Maps



Viewing Images (read only mode)

Images are always displayed at the top right, below Record Address and UUID, if shown

This button will display the image full-screen. Press the spacebar or {Escape} to close the full-screen image.



Click this button to open the image in its own window.

This button restores a zoomed or minimised image to best fit the image window

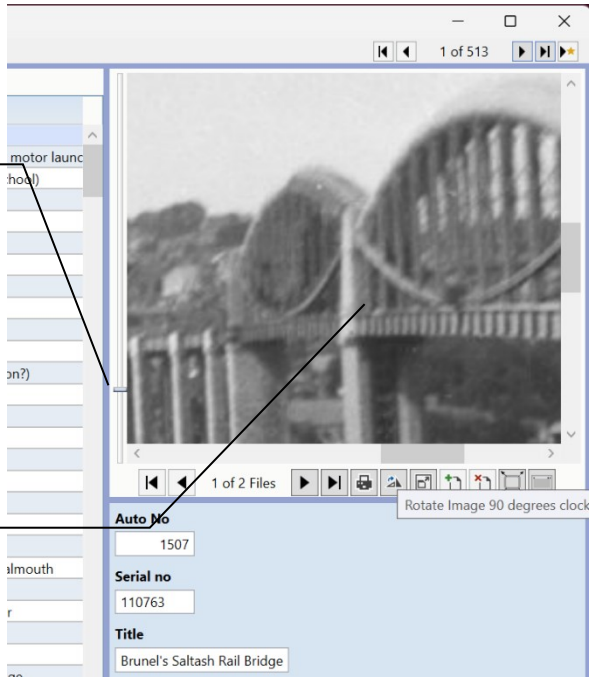
Each record may contain more than one image file. Navigating image files is similar to record navigation, described above. When a record has more than one image file, typically they are the same image scanned at different resolutions

This button rotates an image 90° clockwise.
It is possible to save the rotation

Panning and Zooming images:

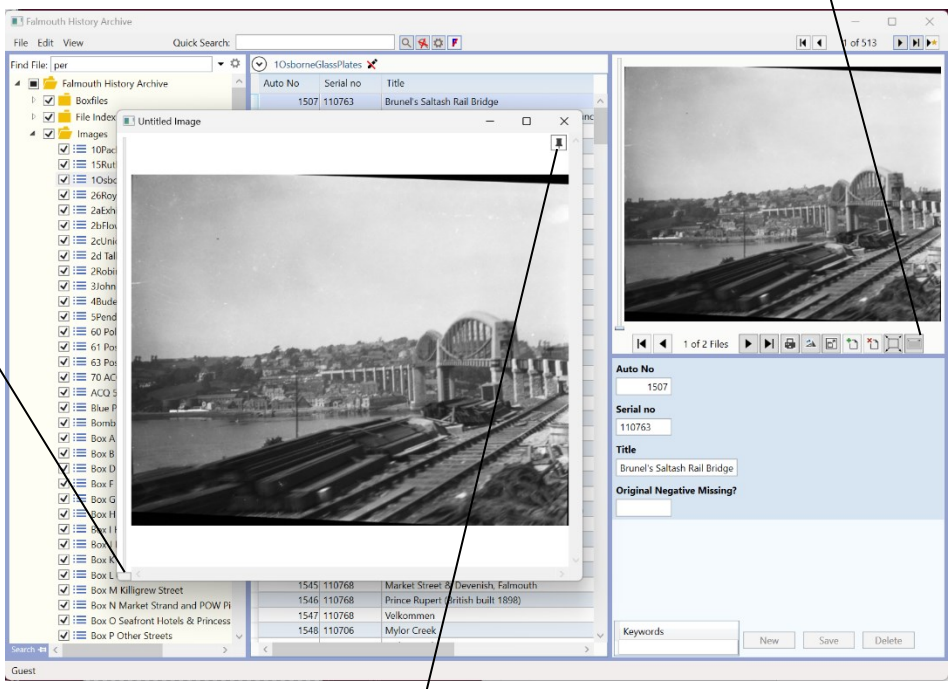
Images can be zoomed by moving the slide bar (at the left of the image) or by using the mouse wheel when the mouse is over the image.
Zoom with the mouse wheel keeps the centre of the zoom under the mouse pointer

Zoomed image can be panned up and down and side to side by dragging with the mouse. You can also use the scrollbars to the side and bottom of the image



Opening new image Windows

Click this button to open new image window



The image window can be zoomed with the mouse wheel or the slider to the left of the image

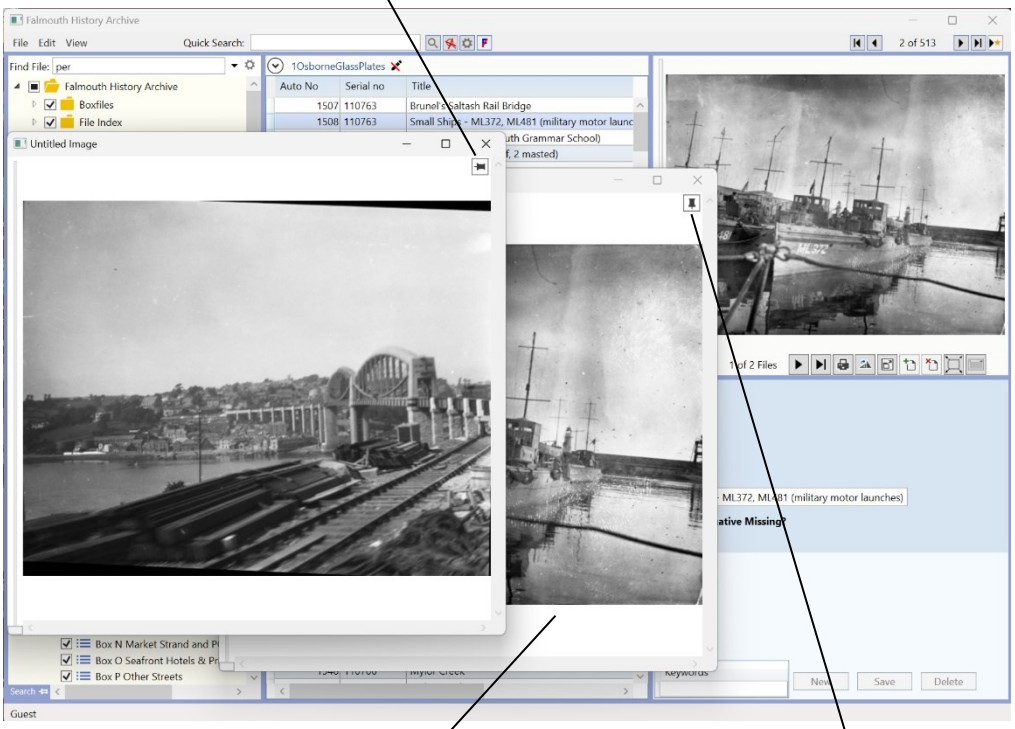
This button toggles the always-on-top window property

The window can be resized as required.

Always – on – top window property

Always - on - top windows sit on top of all other windows, even those of other applications. They remain open even when you move to another record, allowing you to see images from more than one record at the same time.

This is the image from the previous record with the always on top button toggled to keep this window visible...



...this is the image window for the current record, having just been opened. Note that it lies underneath the window from the previous record

Optionally, click the **Always on Top** button for the 2nd window in order to allow it on top of the first

You can open as many image windows as you wish to see side-by-side. Move them around the desktop as needed.

Image windows that are not set *Always on Top* will be closed automatically when you move to another record

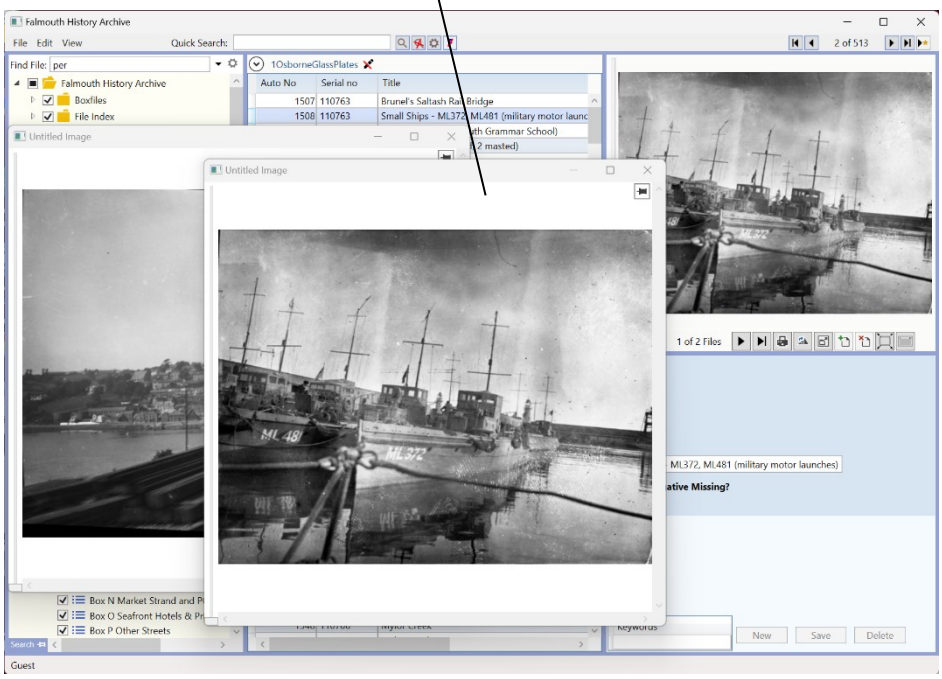


Image Context Menu

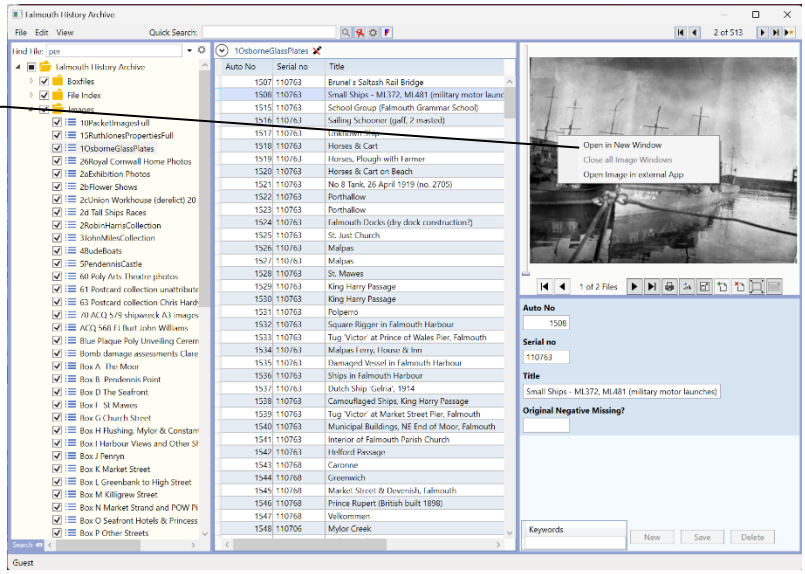
Right-click an image to display the context menu. There are 2 useful commands in read-only mode:

- **Open in a New Window**
- **Close All Image Windows**

The first command is self-explanatory and opens the image in a window of its own, as above. The second command closes all open image windows and is useful when you have several image windows open.
The 3rd command:

- **Open Image in External App**

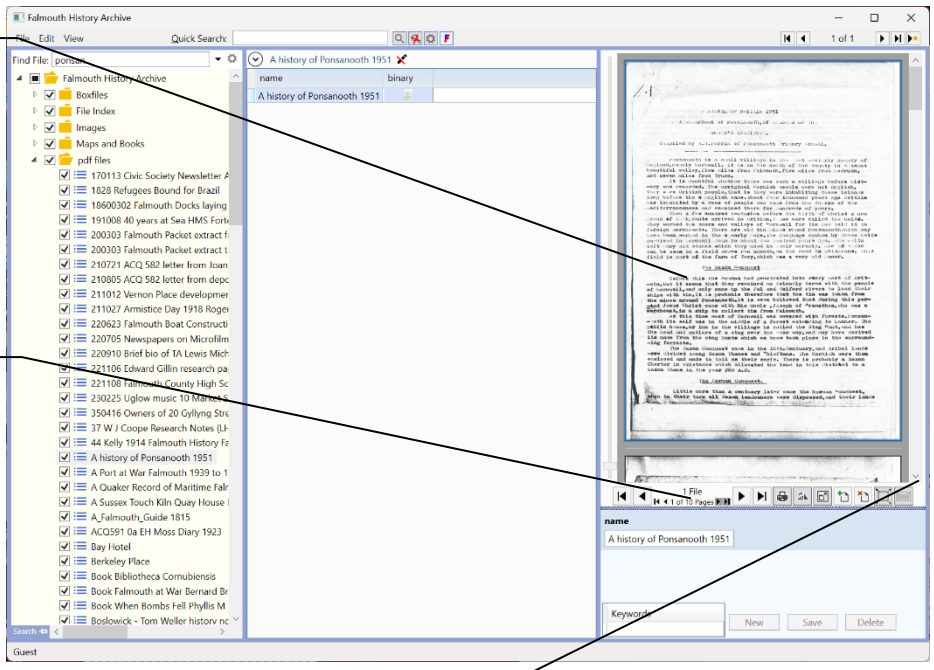
Does not function in read-only mode



Viewing PDF files

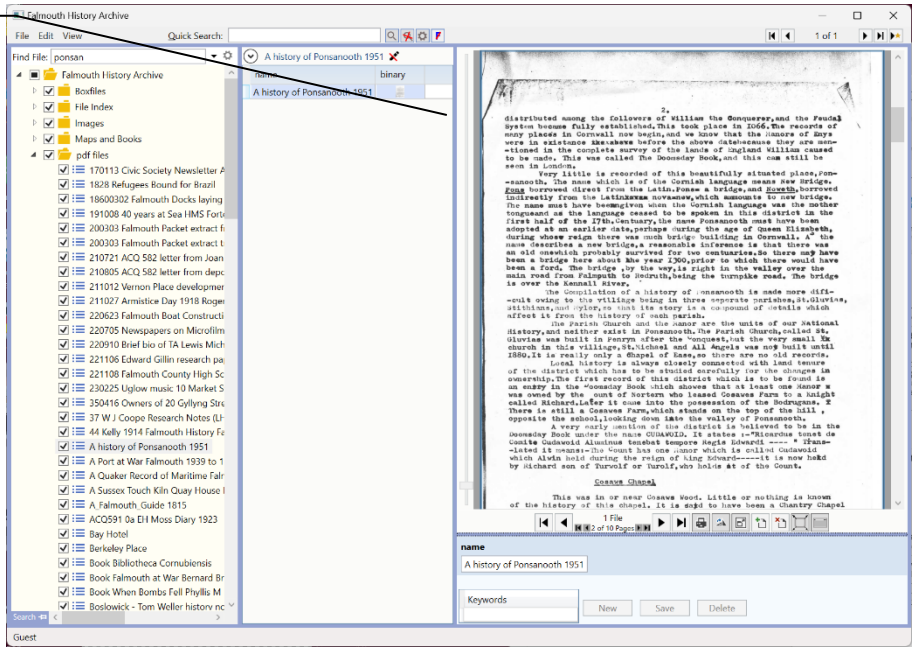
PDF files are displayed in the binary window, in a similar way to images.

Typically, only one PDF file is attached to an individual record but that PDF file may contain several pages.
Use the small buttons alongside the page number to navigate within the PDF file or…



…you can use the scroll bar to navigate between pages

Drag the panel dividers to optimise viewing.
By default, PDFs open sized fit to width. When the vertical panel divide is dragged the PDF will enlarge so that its width continues to best fit the space.



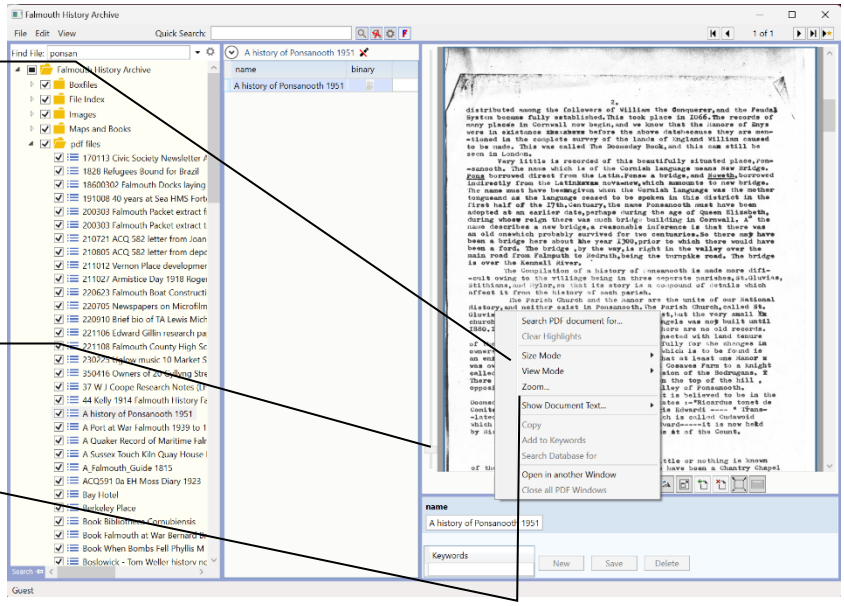
Zooming and panning PDFs

Pan and zoom behaviour of PDF files differs to that of image files.

PDF files have a **Size Mode** and **View Mode**. To change these right-click the PDF to display the context menu.
The default size mode is **Fit to Width** and the default View Mode is **Vertical**.
Size Mode and View Mode have submenus to change the individual properties

Note: the zoom slider is disabled unless the Size Mode is set to **Zoom**…

…a shortcut to change the size mode to Zoom is to click **Zoom** on the context menu



Size mode

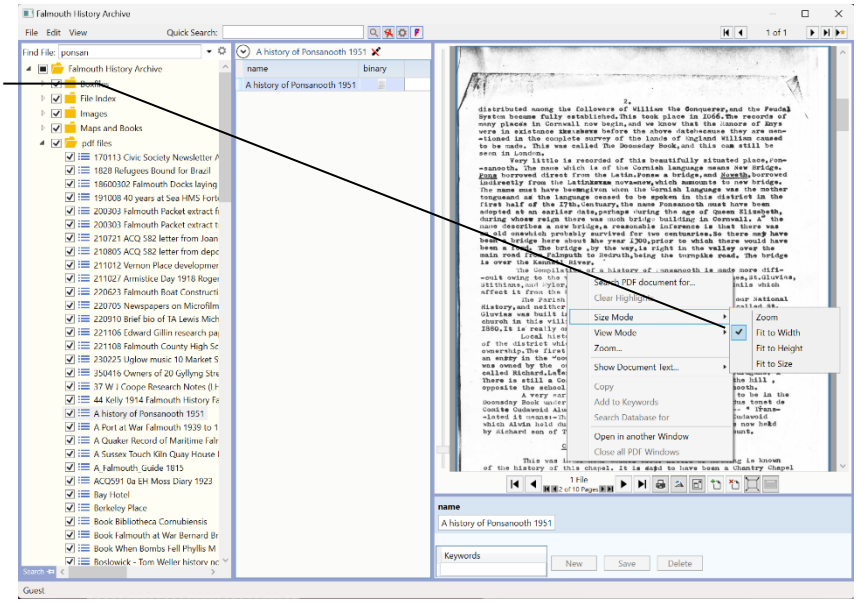
There are 4 size modes:

Zoom – this setting enables the zoom slider to the left of the PDF window. When set to Zoom use the slider to change the display size.

Fit to Width displays the PDF so that its current page width just fits within the PDF window

Fit to Height displays the full height of the current page within the PDF window

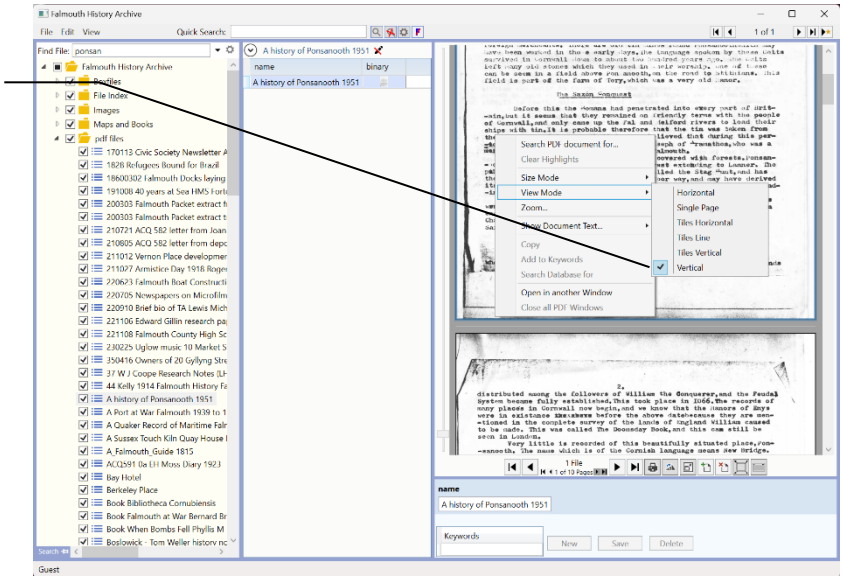
Fit to Size ensures that the whole of the current page is visible within the PDF window, best fitting to both height and width.



View mode

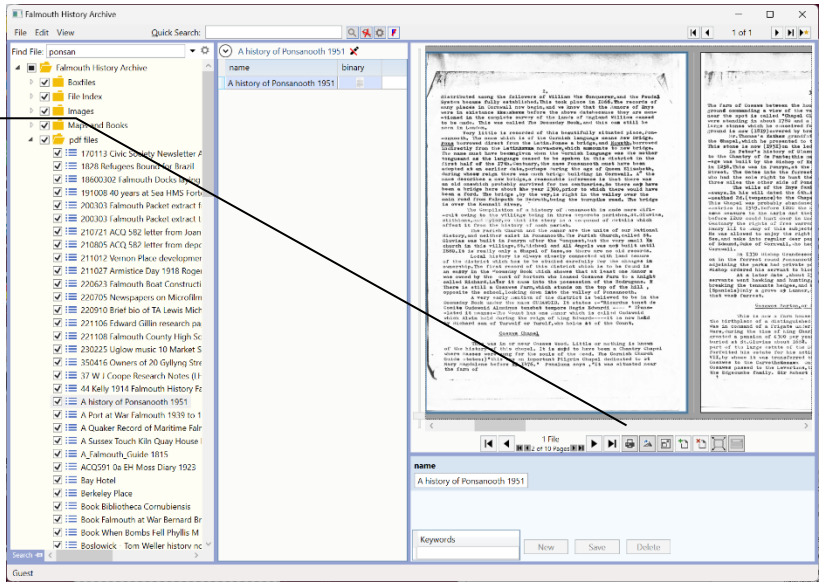
Vertical

The default View Mode is **Vertical**
in this example the Size Mode has been set to Fit to Size.
In Vertical View Mode consecutive pages are arranged vertically. Use the vertical scroll bar to navigate between pages. The mouse wheel scrolls the vertical scroll bar and so can be used to navigate between pages



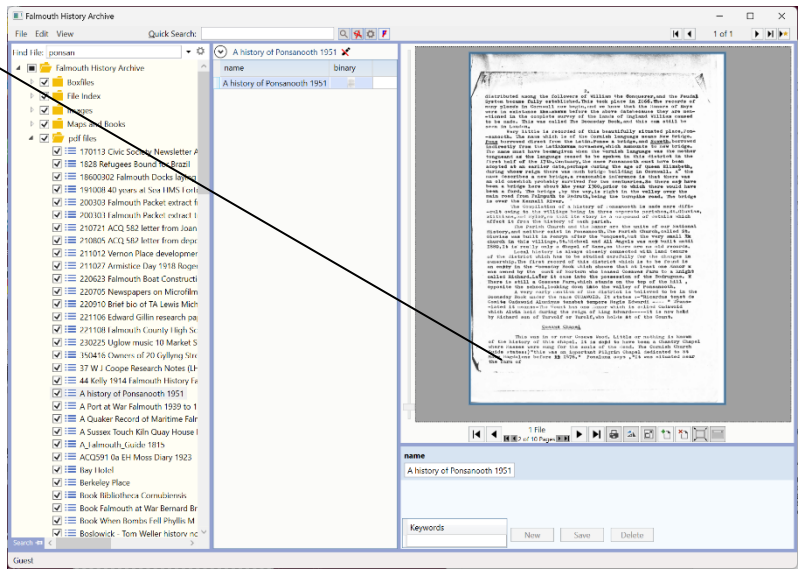
Horizontal

In **Horizontal** view consecutive pages sit side-by-side. As in the previous screen print, the size mode here is set to Fit to Size. The appearance and behaviour of the vertical and horizontal scroll bars depends upon the Size Mode. In this example the horizontal scroll bar has appeared and the vertical scroll bar disappeared. The mouse wheel will, therefore, scroll the document horizontally and navigate between pages. If the size mode is set so that a vertical scroll bar is present the mouse wheel will adjust the vertical scroll bar instead of the horizontal and therefore not navigate between pages.



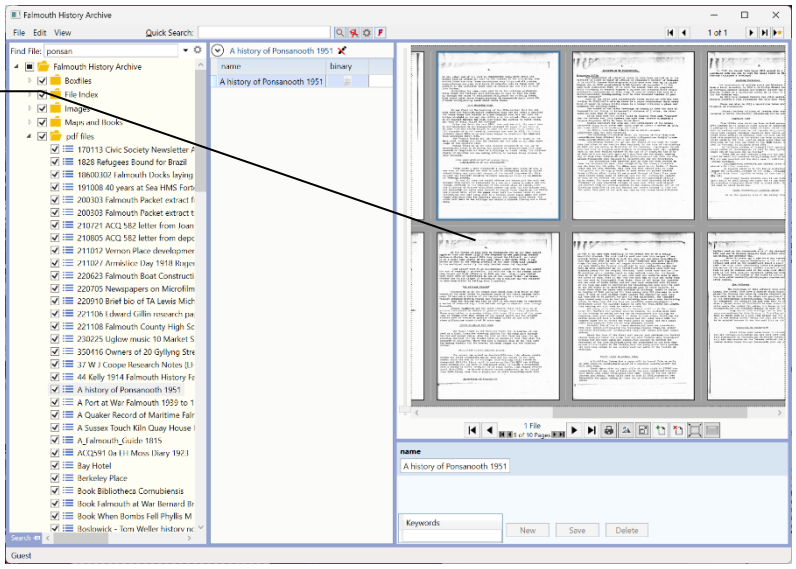
Single page

Single Page View Mode displays only the current page. You cannot navigate between pages using the scroll bars, which have disappeared, you must use the **Page buttons** or type a page number into the **Page Text Box** between the buttons.



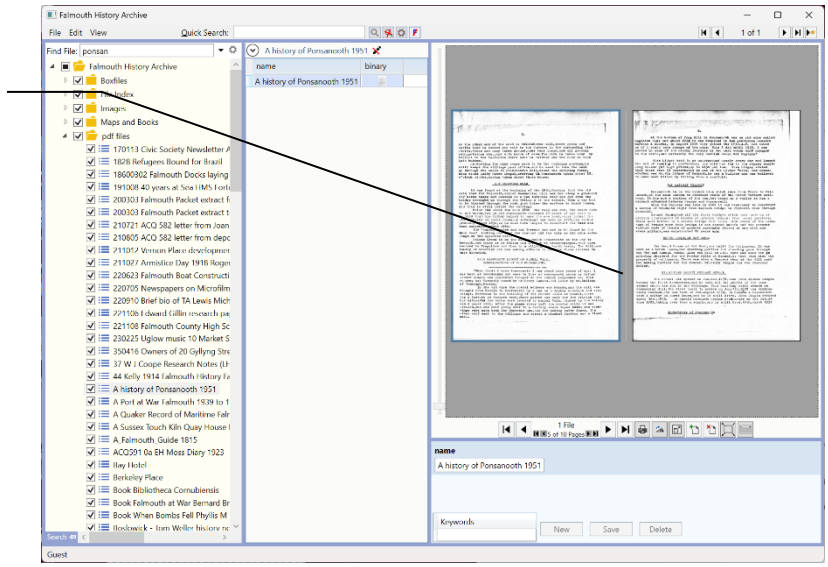
Tiles horizontal

Multiple pages are displayed. They are arranged in 2 rows and ordered with odd numbered pages on the top row and evens on the bottom, page numbers increase from left to right. The horizontal scroll bar will navigate to the end of the document.



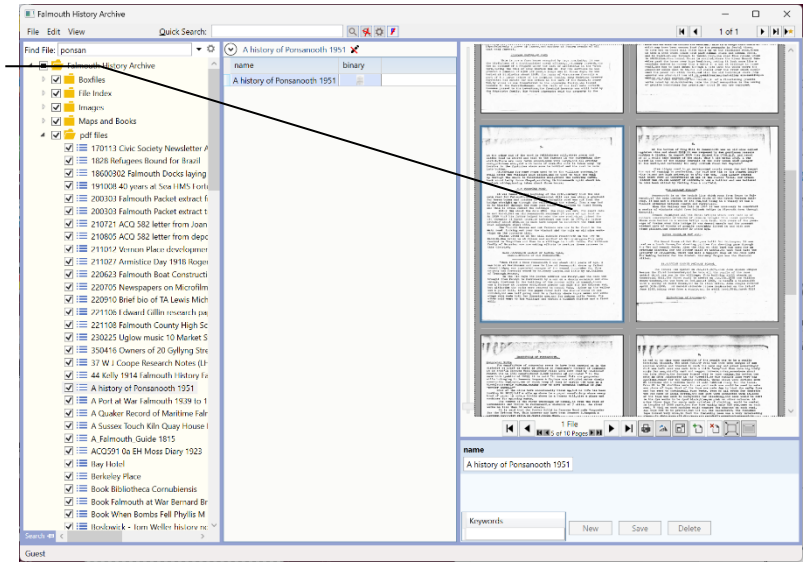
Tiles line

This view displays 2 pages side-by-side with a vertical line between the 2. Use the page buttons to navigate.



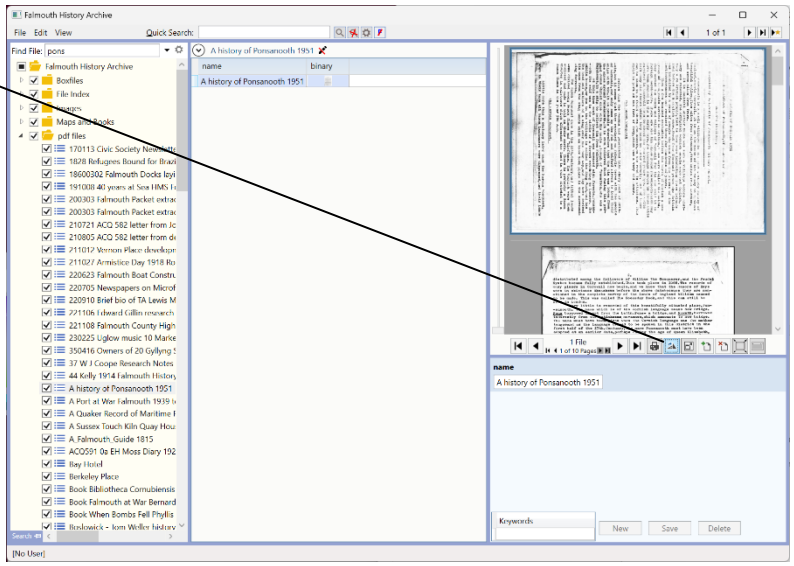
Tiles vertical

Pages are arranged vertically with odd-numbered pages in the left column. Use the vertical scroll bar to navigate to the end of the document.



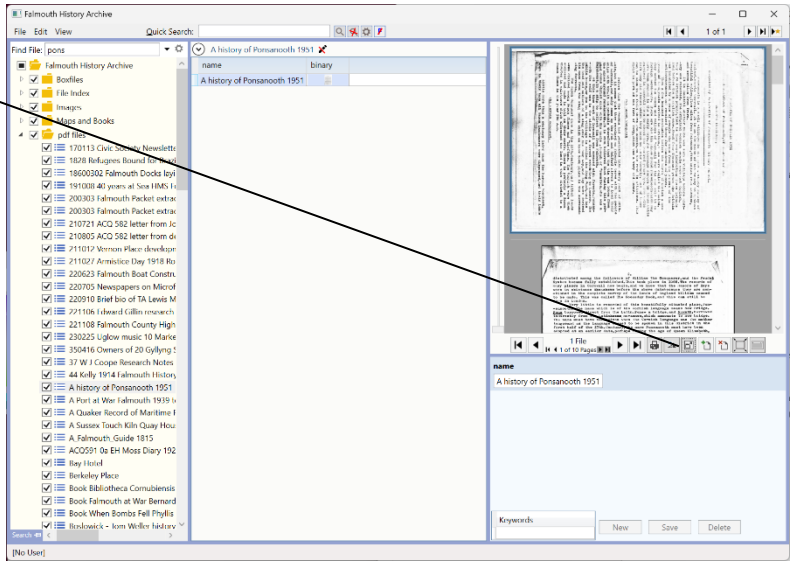
Rotating a page

To change the orientation of the current page, click the **Rotate** button. This will progressively rotate the current page by 90° in a clockwise direction. Orientation changes cannot be saved for PDF files.



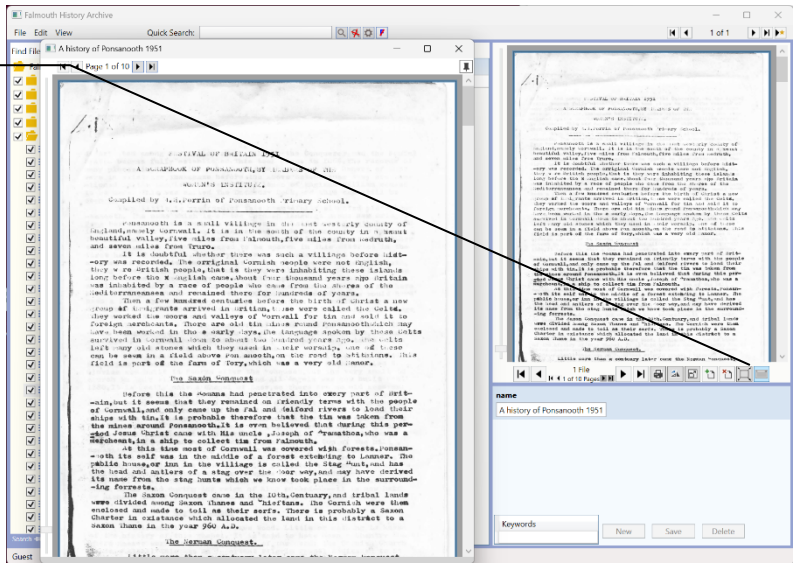
Fitting current page to size

The **Fit to Size** button performs the same task as the **Size Mode**=> **Fit to Size** command on the context menu



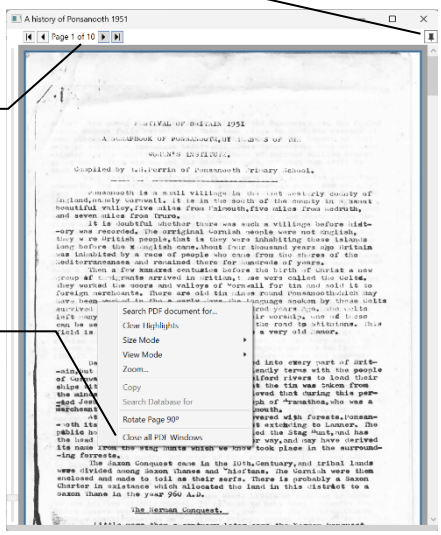
Open PDF in a new window

These 2 buttons both open the current PDF file in a new window.
The open file Fullscreen button (to the left) maximises the new window. This behaviour differs to the Fullscreen button for images which opens a true Fullscreen image. The PDF file version opens a normal bordered window. It can be closed using the cross, top right, or by pressing {Escape} or {Space} on the keyboard.



Like its imaging equivalent, the PDF window can be toggled to sit on top of all other windows. See the image window **Always -On-Top** property, above.

Navigation, view modes etc. are similar to those described above



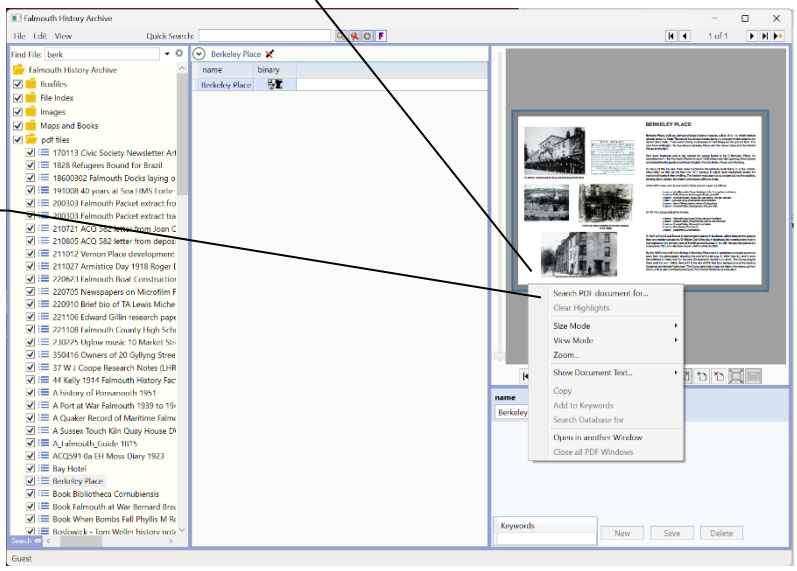
The PDF window right-click context menu has a **Close All PDF Windows** command that will close all PDF windows, whether or not they are set to be **Always -On-Top**. PDF windows not set to be **Always -On-Top** close automatically when you navigate to a different record.

Searching PDF documents

The text of PDF documents within the database can be word-searched globally using the **Quick Search**. Individual PDF documents can be searched in more detail by using the **PDF document search** window accessed by using the right click context menu. The PDF document search allows searching for phrases, partial words and case matching.

Right click the PDF to display the context menu…

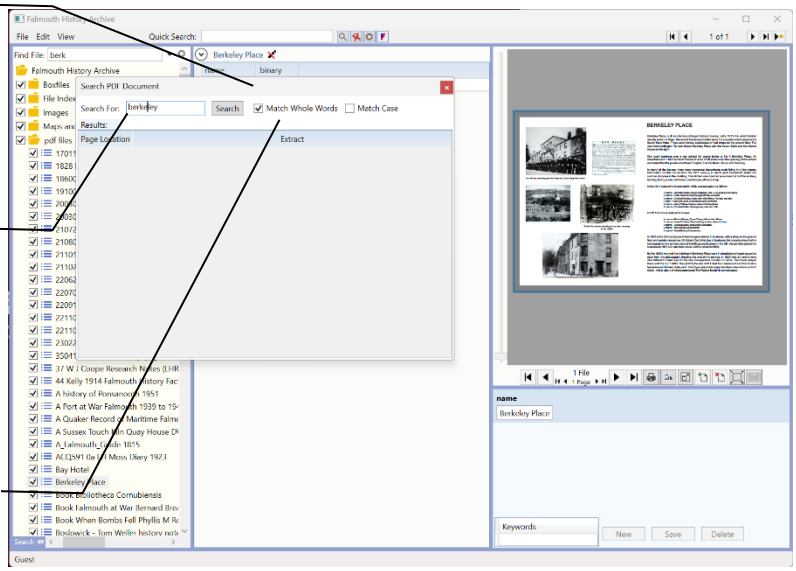
…click **Search PDF Document for**…



…this will open a search window that sits on top of all other windows

Enter search criteria here

There are 2 checkboxes that allow you to match whole words or match case



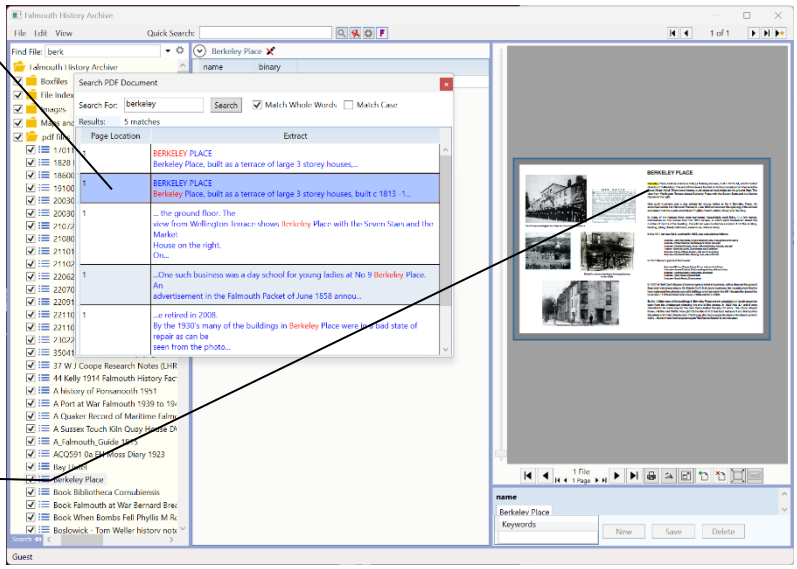
The **Search** button begins the search

The number in this column is the document page number containing the match

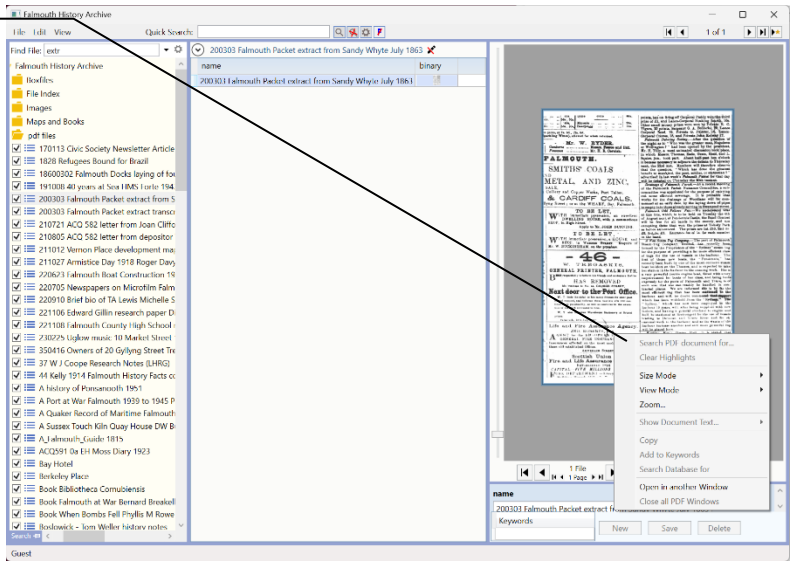
Search matches are highlighted in red

Click a row…

…to highlight the match in the PDF document

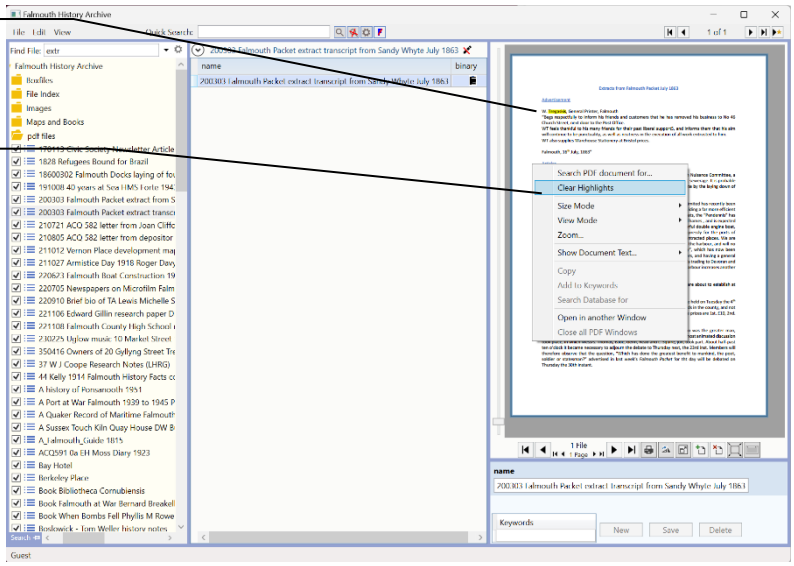


Some PDF files are simply an image of the document text (that is they contain no editable or searchable text). In that event the search command is disabled.



Text highlights in the document can be cleared by…

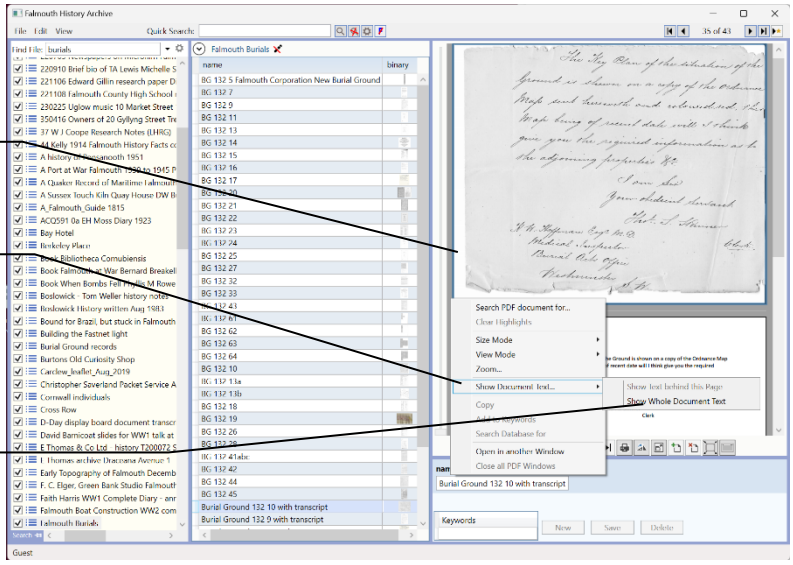
…right clicking to display the context menu and choosing **Clear Highlights**



View the PDF document text

Right click the document to display the context menu…

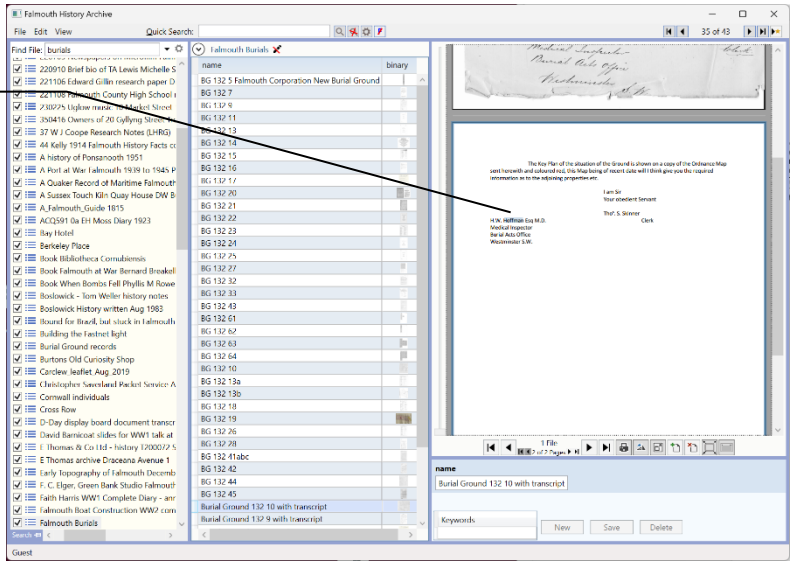
…click **Show Document Text** to display the submenu…



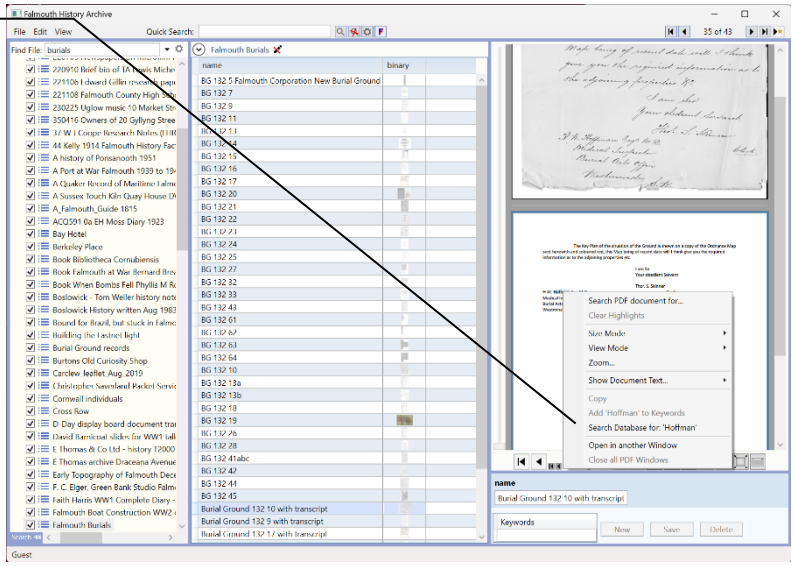
…the submenu has 2 commands. These allow you to **View the Text Behind the Current Page** or the **Whole Document**. In this example the current page is an image of handwriting and has no editable text behind it. The command to view text for the current page is therefore disabled.

Searching the database from text within a PDF

Editable text within a PDF can be selected with the mouse. You can double-click a word or drag through the text. Having selected some text, right click to display the context menu…

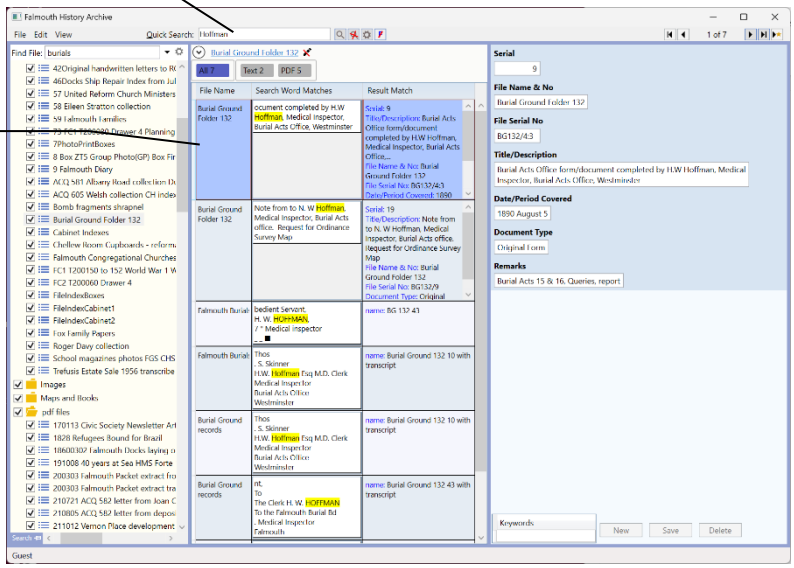


…click on the context menu command **Search Database For 'xyz'**…

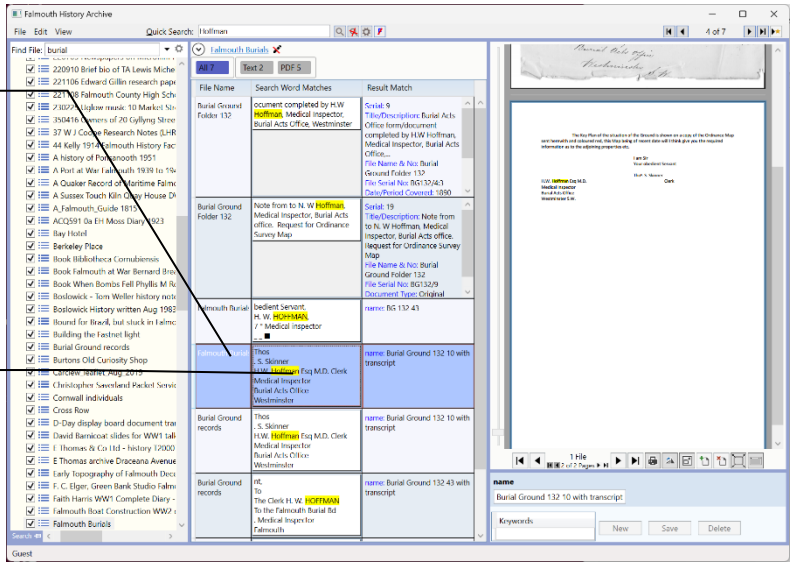


… This will place the selected text in the quick search criteria box and run the search…

…the results are shown here

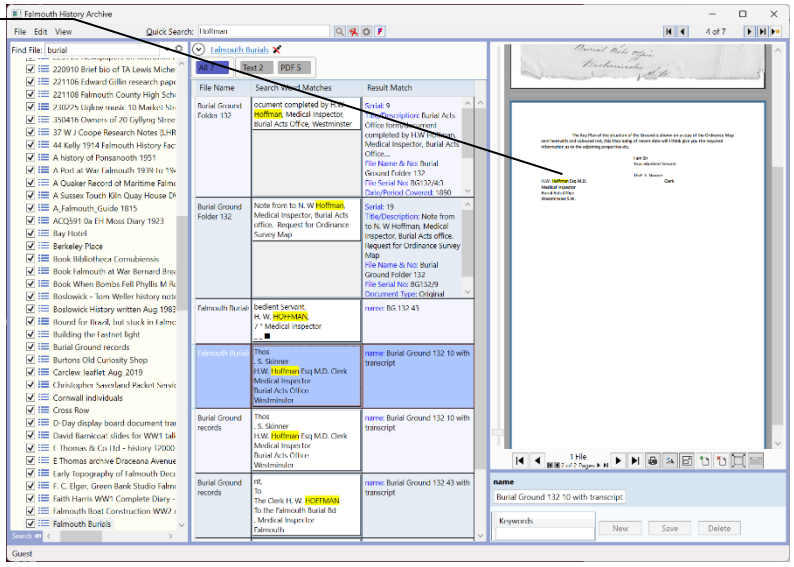


In this example I have clicked the record that corresponds to the PDF from which we initiated the search. This has displayed the PDF once more…



…the match is highlighted in yellow in this grid cell. Click this grid cell…

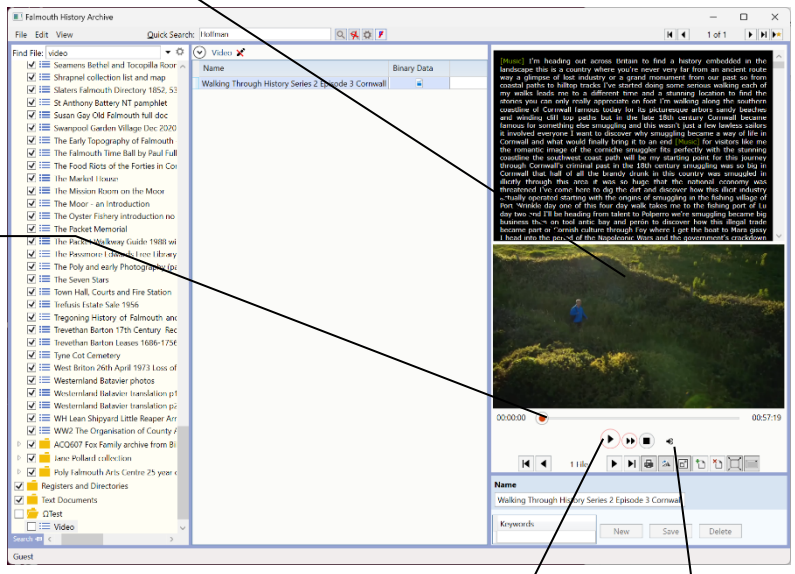
…to navigate to the relevant PDF page and highlight the match.



Multimedia files

ArchiveDb can play audio and video files and display a searchable audio transcript.

This is an example of a video with an audio transcript. The video is displayed below the transcript.
You may need to enlarge the panel by dragging the divider to see the transcript.

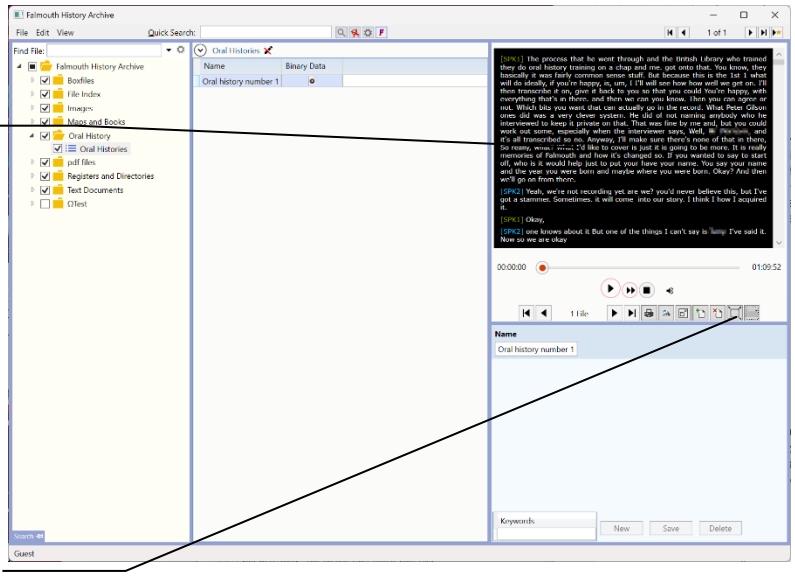


Beneath the video is a slider that control the location within the file.
Drag the **Seek To** slider to move through the multimedia file.
The position within the multimedia file and the total duration of the multimedia file are displayed respectively at the end of the slider.

These 4 buttons respectively **Play/Pause**, **Fast Forward** and **Stop** the video or audio file.
Press the **Fast Forward** button repeatedly to double the fast-forward speed.
Pause stops the multimedia file at the current location. It can be restarted by clicking the **Play** button.
The **Stop** button stops the multimedia file and returns to the beginning of the file.

This button toggles the display of the volume slider:

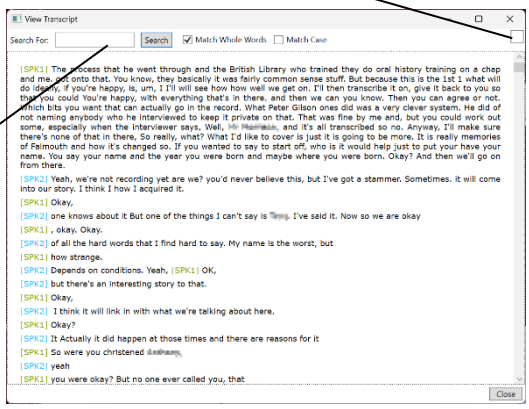
This is an example of an audio file in an oral history record. The transcript is displayed here. Beneath it are the same controls as with the video file, above.



The **Open File in a New Window** button opens the transcript in a separate window…

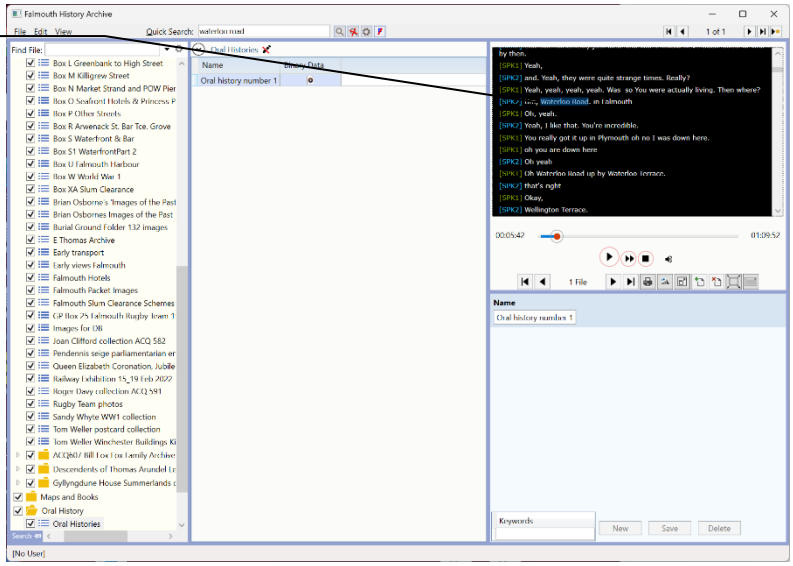
The transcript window has the **Always-on-Top** toggle button common to the image and PDF file viewers

The transcript window also has a text search. The checkboxes can be checked to specify matching whole words or matching case (capitalisation). Search matches are highlighted in yellow.



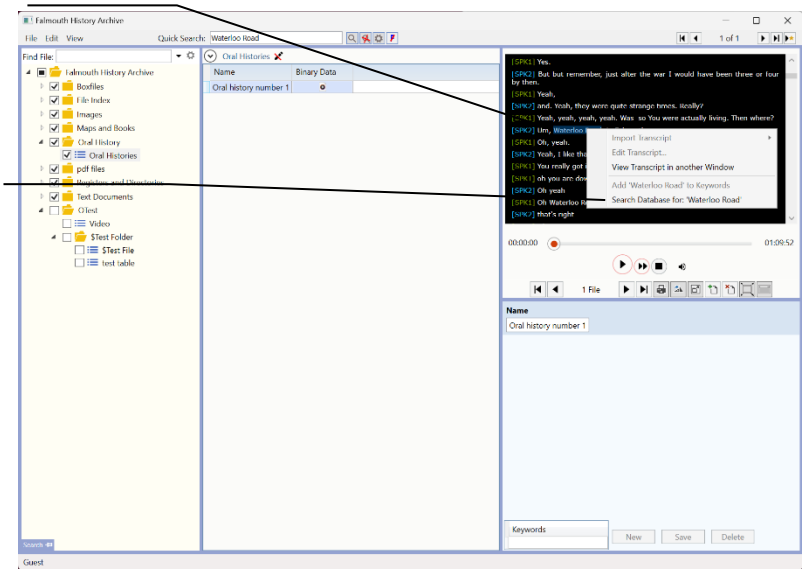
Searching the database from a multimedia transcript

Select the text you wish to search for by dragging through with the mouse. In this case we have highlighted "Waterloo Road"…



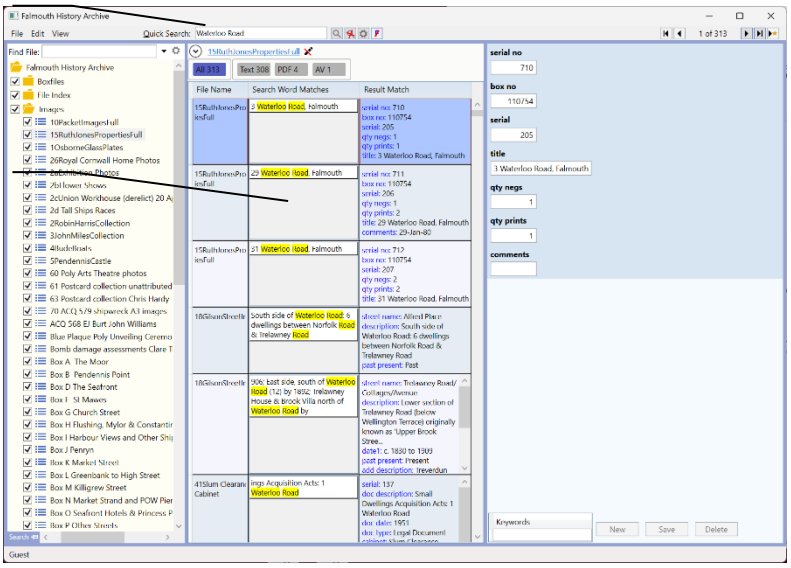
…right click to display the context menu…

…click **Search Database for** ‘xyz’…



…and the search text will be entered in the **Quick Search** text box…

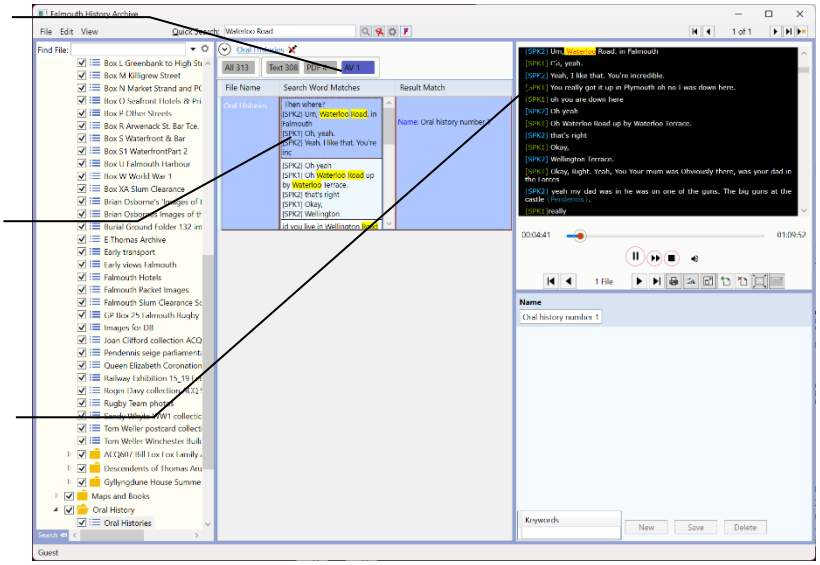
…with the results of the search displayed in the Result Grid



To filter multimedia results in the result grid you can click on the '*Other*' filter button

The **Search Word Matches** grid displays a cell for each match within the transcript. Click on one of the cells to highlight the match…

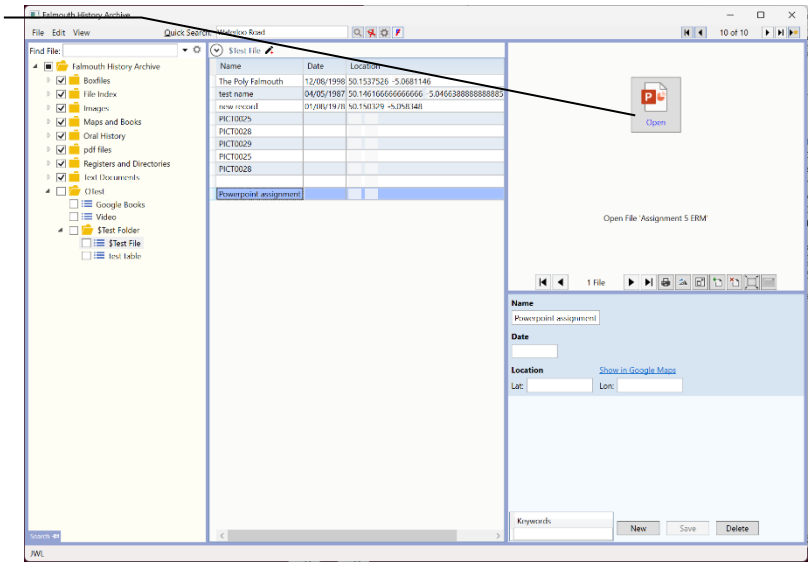
…in the transcript. The audio or video file will play for a few seconds before and after the word match.



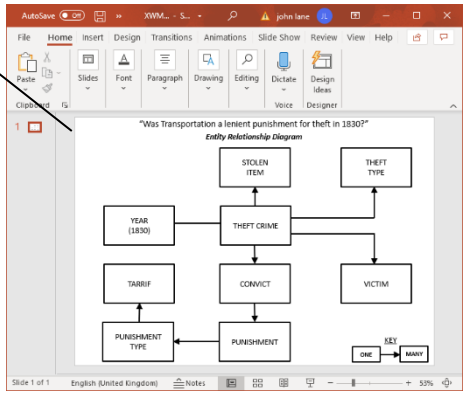
Other Binary Files

ArchiveDb can store any type of computer file but is unable to natively search, index or display the contents of files other than PDFs, text files (rich and plain) and audio transcripts. Instead of natively displaying files within the ArchiveDb user interface, other binary files are displayed within their associated application. This means that, for example, a PowerPoint presentation stored within ArchiveDb would launch Microsoft Office or an alternative programme that had been installed (e.g. OpenOffice) on the relevant PC for viewing PowerPoint presentations.

In this example a PowerPoint presentation has been attached to this record. The binary pane displays a file icon with the word '*Open*', beneath. Click the icon to open the file…

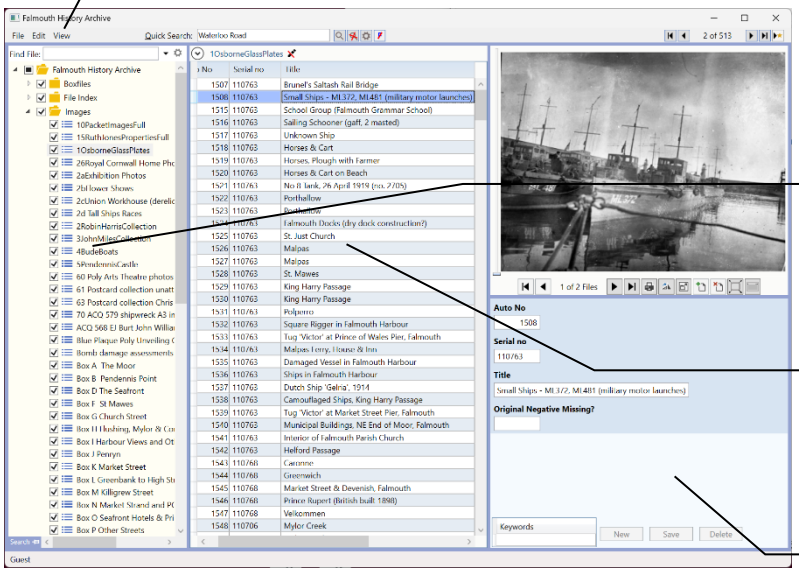


…this launches PowerPoint and opens the associated file.
The contents of this file cannot be searched from within ArchiveDb.



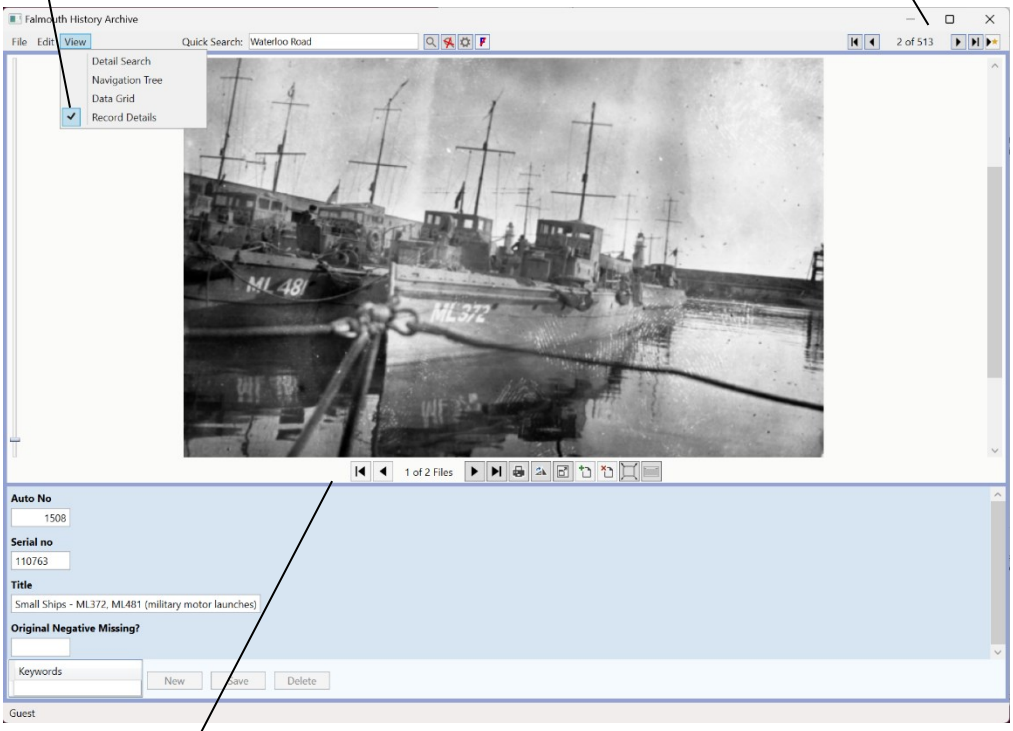
Customising the workspace: View menu

Open the **View** menu and click items to toggle their visibility



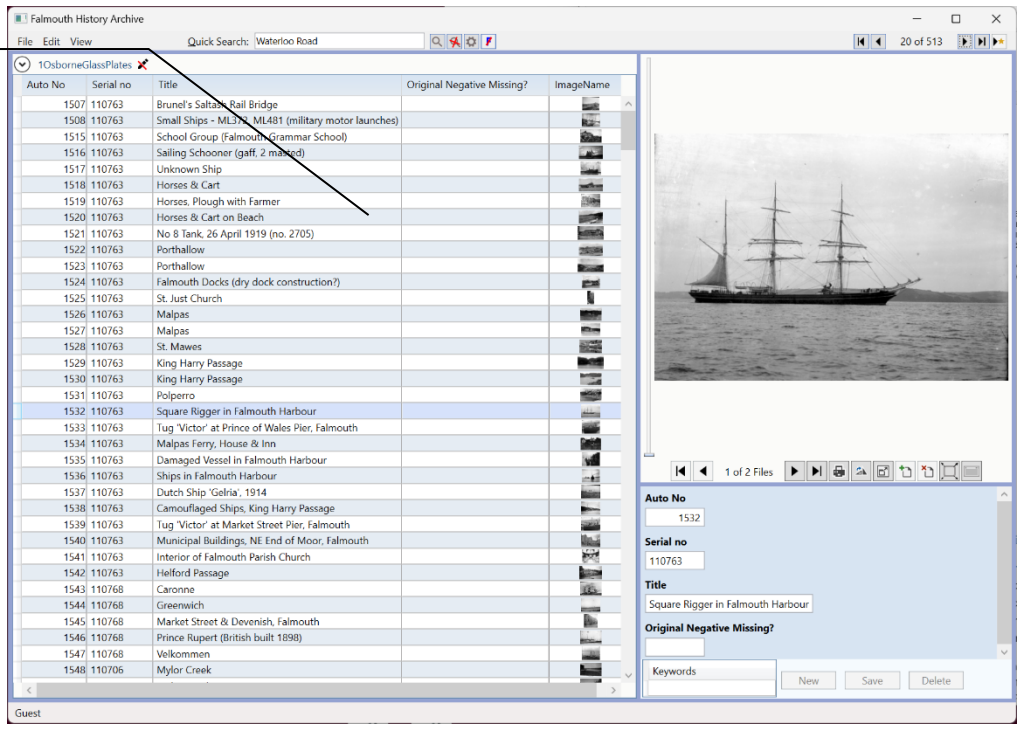
This view has only the **Record Details** visible

Navigate between records using the record navigator



Individual panes can be resized by dragging the dividers, in this case to maximise the space available for the image.

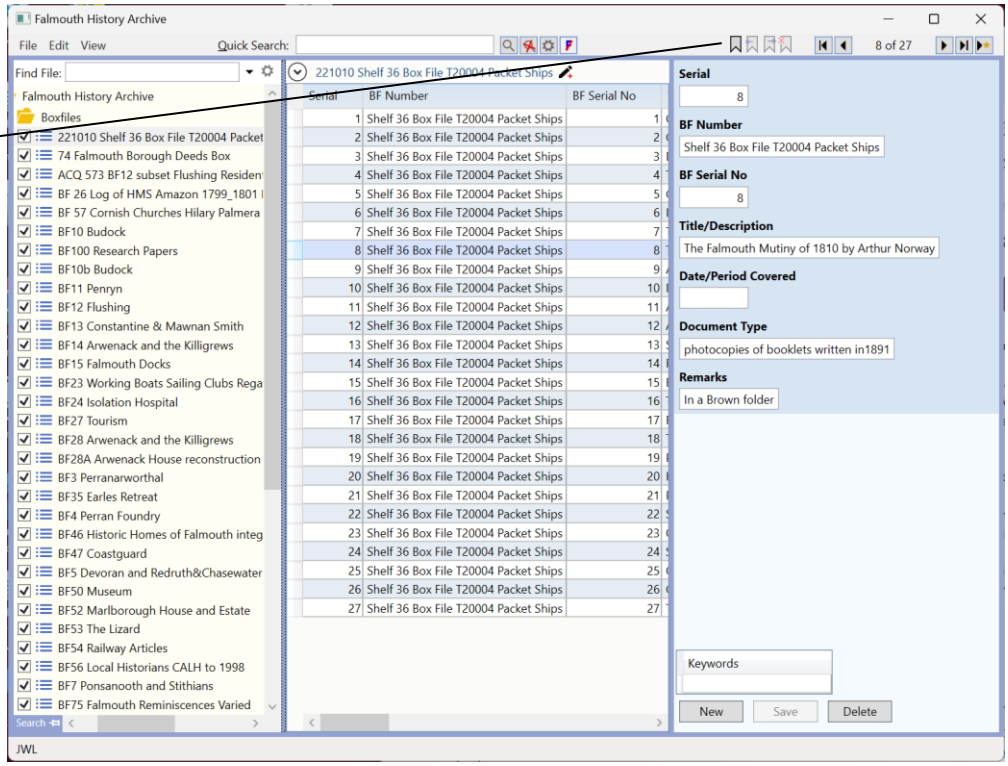
In this example the Navigation Tree has been hidden, just displaying the **Data Grid** and **Record Details.**
This divider has been dragged to optimise the view of the data grid.



Bookmarks

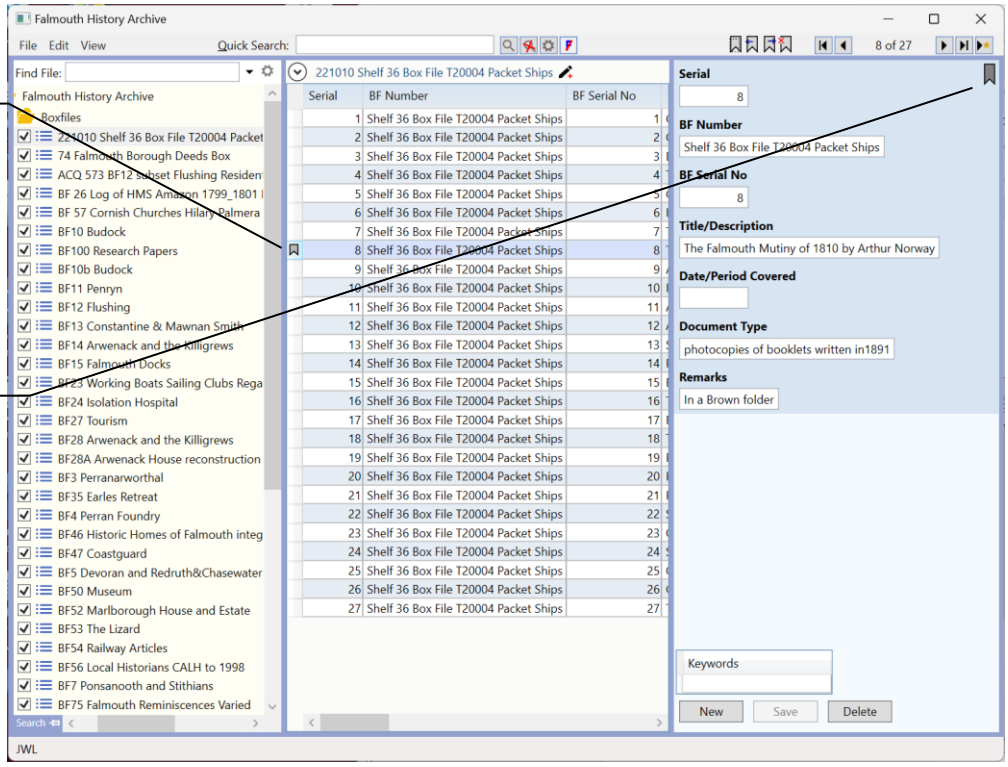
Bookmarks belong to you, the current user, and are not shared with, or visible to, anyone else. They are place markers that allow you to return to specified records. You might, for example, choose to add a bookmark to records found in a search that you wish to return to later.

The bookmark buttons are to the left of the record navigation buttons. The leftmost button toggles a bookmark on and off for the current record.

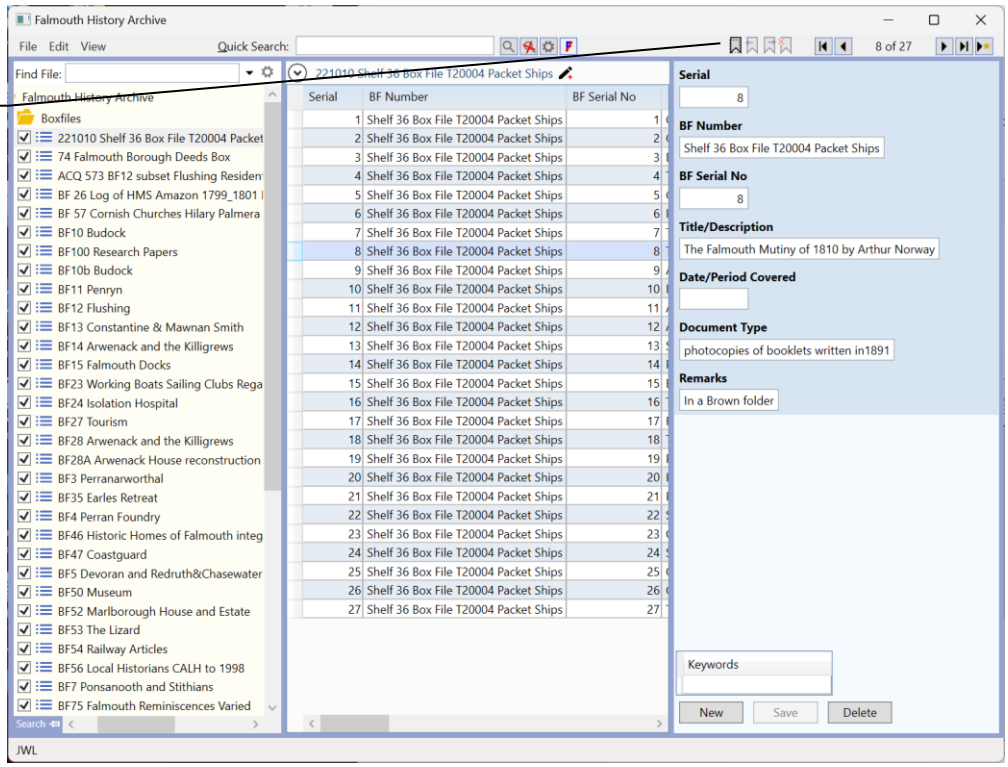


After clicking the button and adding a bookmark to this record, a bookmark appears in the record header in the datagrid…

...and top right in the record details



Click the button again to clear the bookmark for the current record.



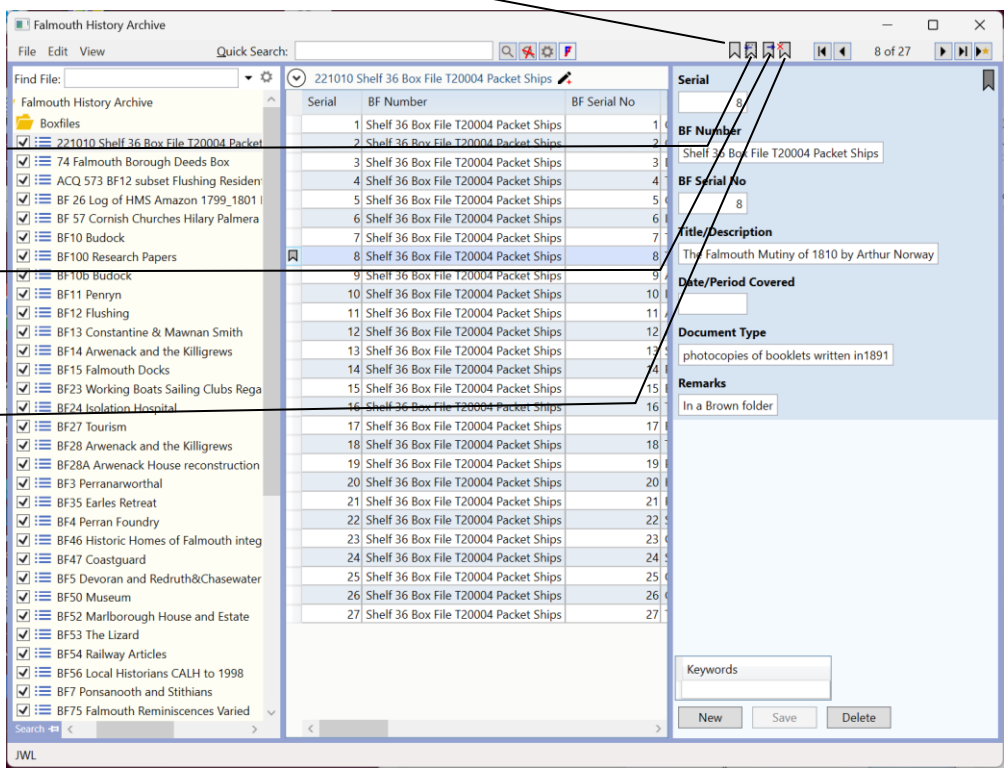
Toggle bookmark on/off for current record.

Navigate to previous bookmark. This will cycle backwards through each bookmark.

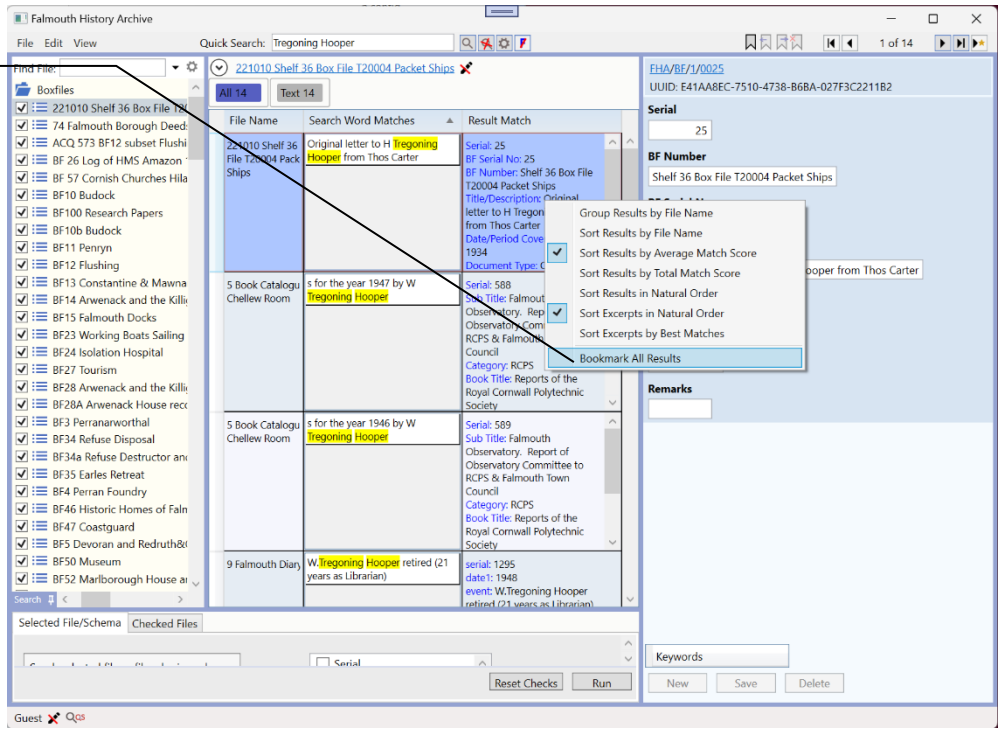
Navigate to next bookmark. This will cycle forwards through each bookmark.

Clear all bookmarks.

The tooltip that will appear when you hover the mouse over this button indicates how many bookmarks you have.

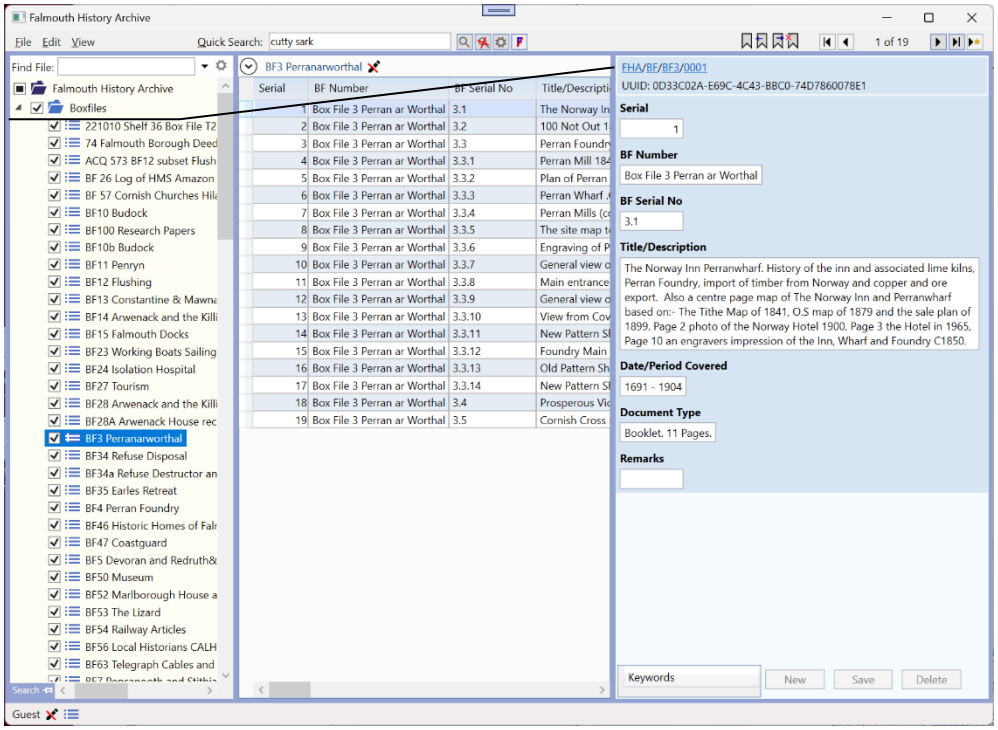


For the **Quick Search** (only, for now) the context menu has a **Bookmark all Results** command.



Hierarchy Reference number

The **Hierarchy Reference number** is shown at the top of the **Record Details Pane**. Its visibility can be turned off; therefore it may not be visible.

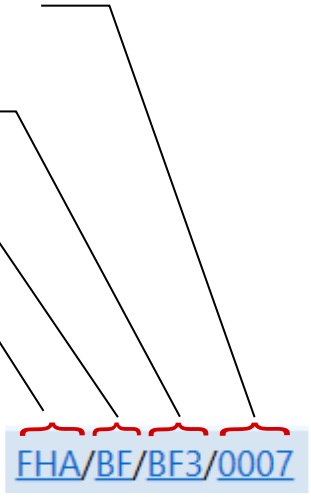


The reference number ends with the selected object, in this case a record with serial number 7.

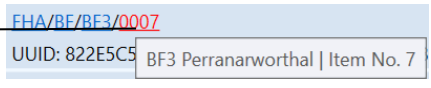
The parent object reference is shown to the immediate left, in this case the object is the **File** *BF3 Perranarworthal* which has reference code ***BF3***

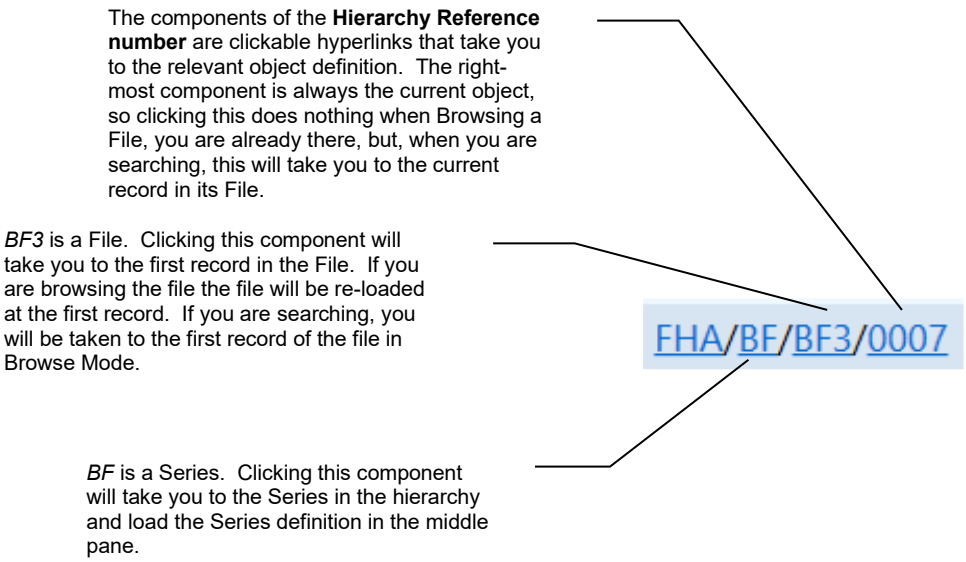
The parent of File *BF3 Perranarworthal* is **Series** *Boxfiles* with reference code ***BF***.

The leftmost reference is the root of the hierarchy. In this case *Falmouth History Archive* with reference code ***FHA***



If you hover the mouse over the **Hierarchy Reference number,** the object title is displayed as a tooltip. In this example the reference is to a record and the tooltip shows the File Name followed by the record serial number.





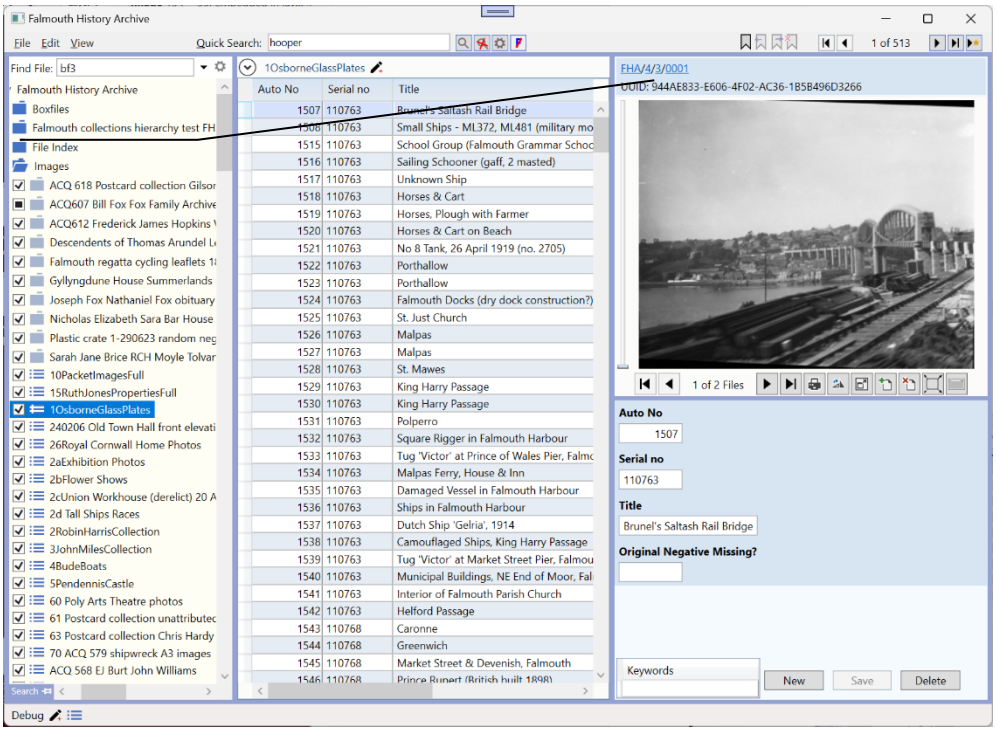
Universally Unique Identifier (UUID)

Each record in ArchiveDb now has a UUID. A UUID is a 128-bit label, typically represented as a 36-character hexadecimal string, eg
6E8DF737-BD7B-445B-8FAF-D8AA06EF0BAC
these are huge numbers.

For practical purposes, UUIDs are considered unique although they are not guaranteed to be so. For example, generating 1 billion UUIDs per second for 100 years would have a 50% chance of a single duplication. For this reason they are often used for identifiers across distributed systems: this means that data assigned a UUID by another database system (it doesn't need to be ArchiveDb) and imported to ArchiveDb will retain its UUID unchanged.

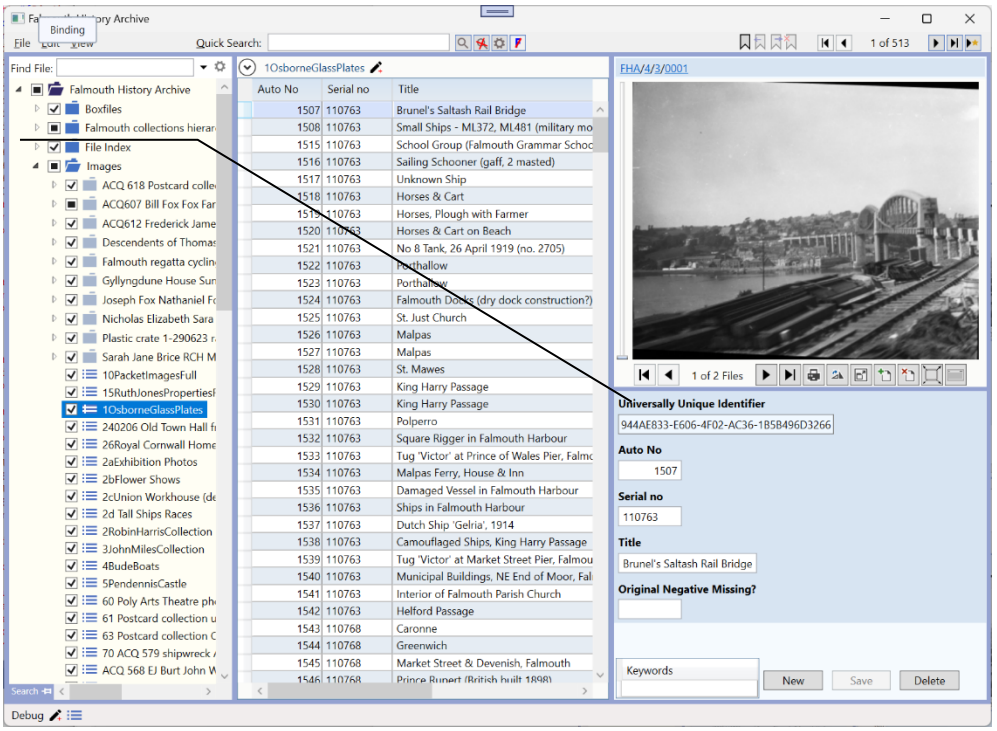
The practical benefit of UUIDs is mainly applicable to cataloguing physical objects. It's not so useful as a unique identifier of a piece of abstract information. With respect to physical objects, the UUID, or a bar code or QR code generated from it, can be used as an object label that refers to the object's database entry. If the object were transferred to another collection that also used UUIDs, the original UUID would/could continue to refer to the object's database reference in the new collection and new database owing to the universally unique nature of UUIDs.

The UUID is optionally shown at the top of the **Record Details Pane**, below the **Hierarchy Reference number**. Its visibility can be turned off; it may, therefore, not be visible.
In this example this UUID would provide an identifier for the physical glass plate whose image is shown here.



In this example UUID visibility has been turned off at system level, but turned on for the schema.

This results in the UUID being moved from the top of the detail pane to displayed as a data field, instead.



Searching

Keywords

Keywords can be set at file level and individual record level. There is no limit to the number of keywords that can be attached to a record or file. Keywords must be single words, not phrases. They can be non-standard words (neologisms), for example words made from short phrases run together without the spaces.

The same keyword might be attached to several files to identify and return certain file groups in a search.

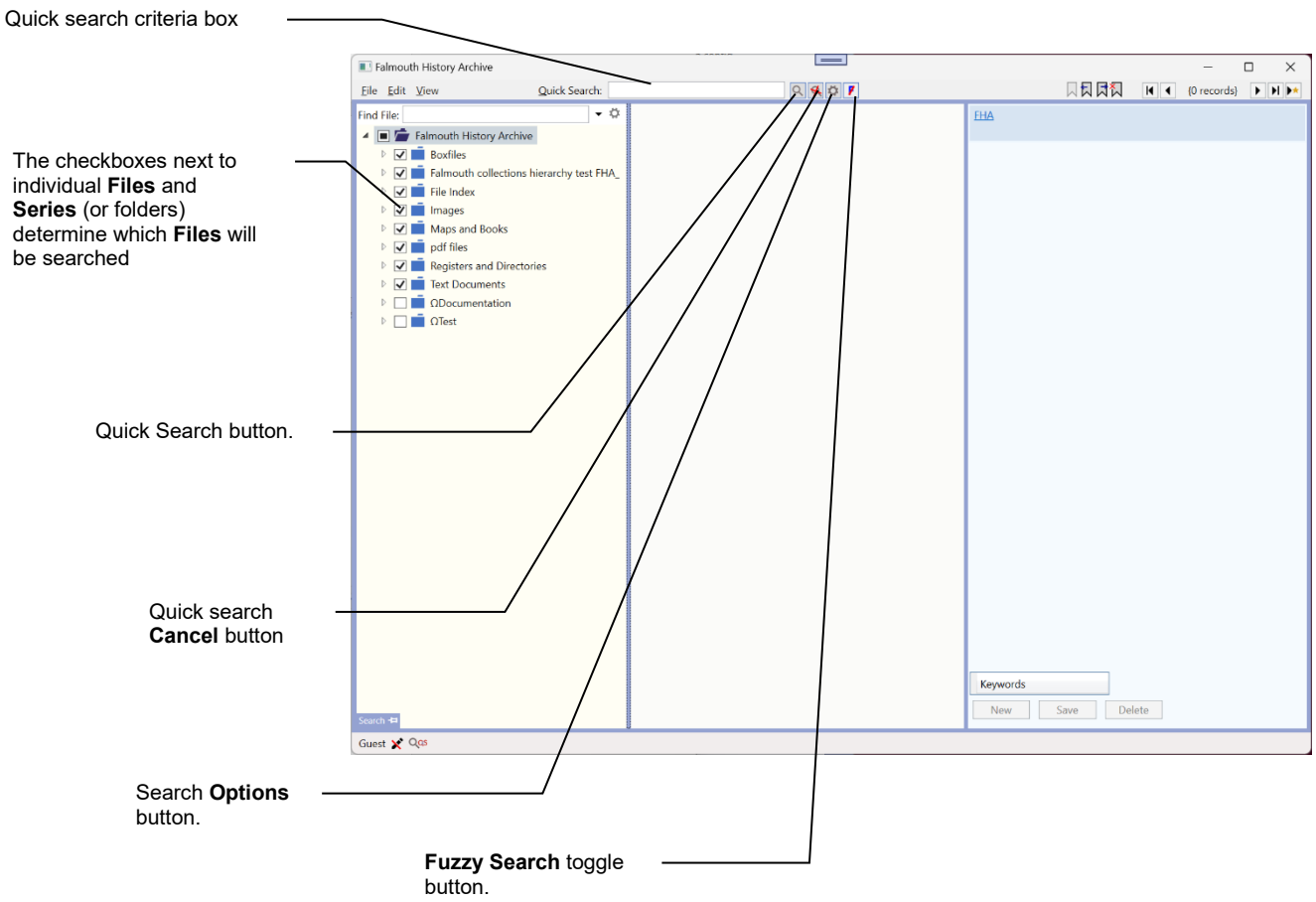
Record indexing and searching

When a record is saved, the record is *parsed* to individual words. Common words are discarded and the remaining words saved to a separate database index. The parsing process also identifies dates within a text field and saves these to a second index. If a record contains a date field[1] that data is also saved to the date index. The process of identifying dates within free text is imperfect and so some information that 'looks' like a date, but isn't, might be saved to the date index, possibly leading to some spurious search results.

[1] See later for description of data types

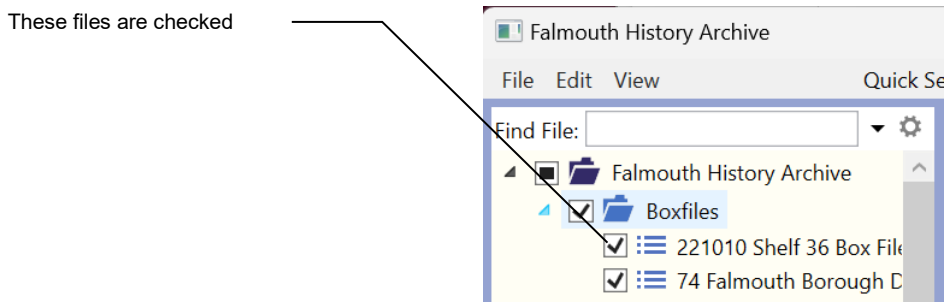
Quick Search

The Quick Search (QS) searches the text content of the checked Files including the text content of imported PDFs and the transcripts of Audio Visual files. It can search for dates as well as text. Results can be bookmarked using the **Bookmark all Results** command on the results context menu.



Checked files

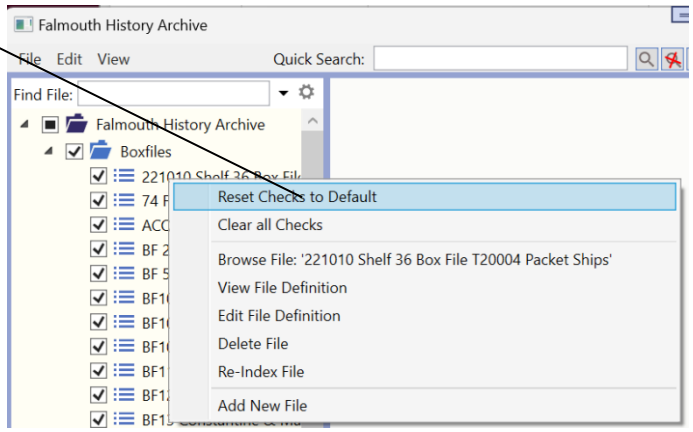
Common to all the searches, QS only searches ***checked files***. Checked files are those with a tick in the checkbox to the left of the file name.



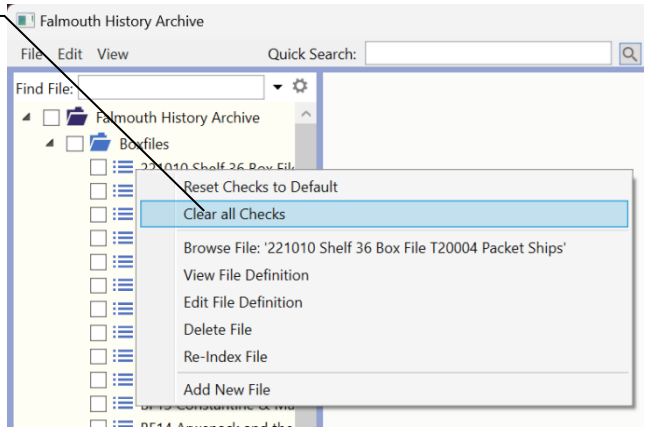
When ArchiveDb loads, or the login is changed, the TreeView loads with the files checked as specified by their default checked setting. If you change the check status of the files you can go back to the default checked status using the context menu command **Reset Checks to Default**. The

training and documentation files are not checked by default and you would usually wish to exclude these from your searches.

Right-click the TreeView and select **Reset Checks to Default** from the context menu

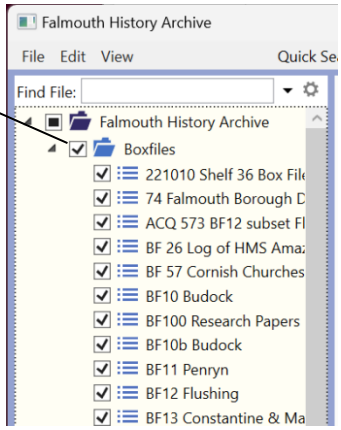


If you wish to check a sub-set of files for a targeted search, you may wish to clear all the checks before checking the ones you require. Right-click the TreeView and select **Clear all Checks** from the context menu to remove all checks

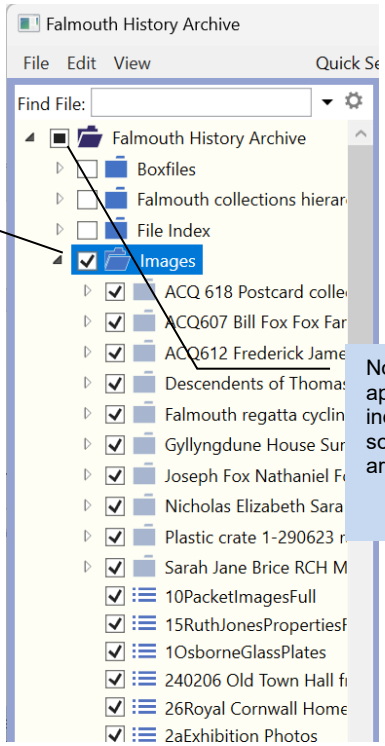


The search selection checkboxes can all be toggled between de-selection and selection by clicking on the checkbox at the root of the tree or root of a branch.

In this example all the child Files in the *Boxfiles* folder become checked when the parent, *Boxfiles*, is checked.



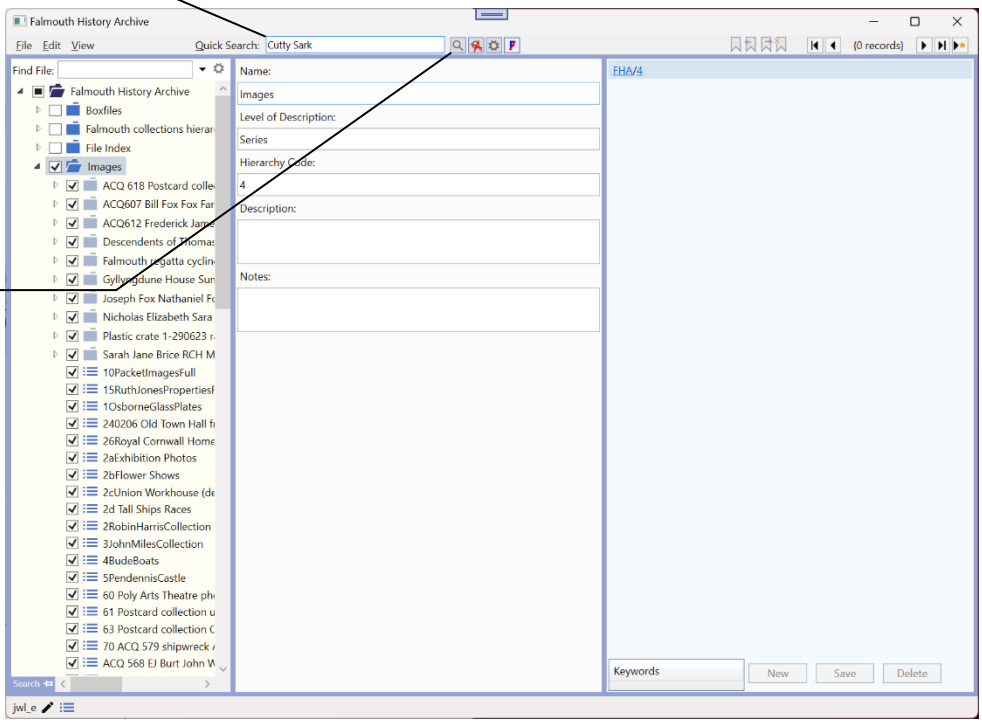
In this example, to search only files within the *Images* series, deselect all the checkboxes using the context menu command, then select all the files within the *Images* series by checking the *Images* node



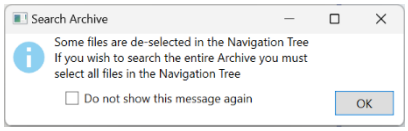
Running the search

In this example I want to look for images of the Cutty Sark.
Begin by typing the words *Cutty* and *Sark* into the criteria text box. To commence the search you can either press {Enter} on the keyboard or click the search button

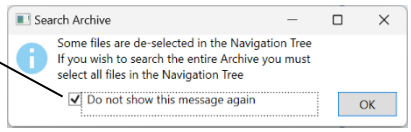
Search button



This message is displayed when some Files' checkboxes are deselected and consequently you are searching a subset of the overall data. It is quite easy to forget that you have deselected some of the Files.



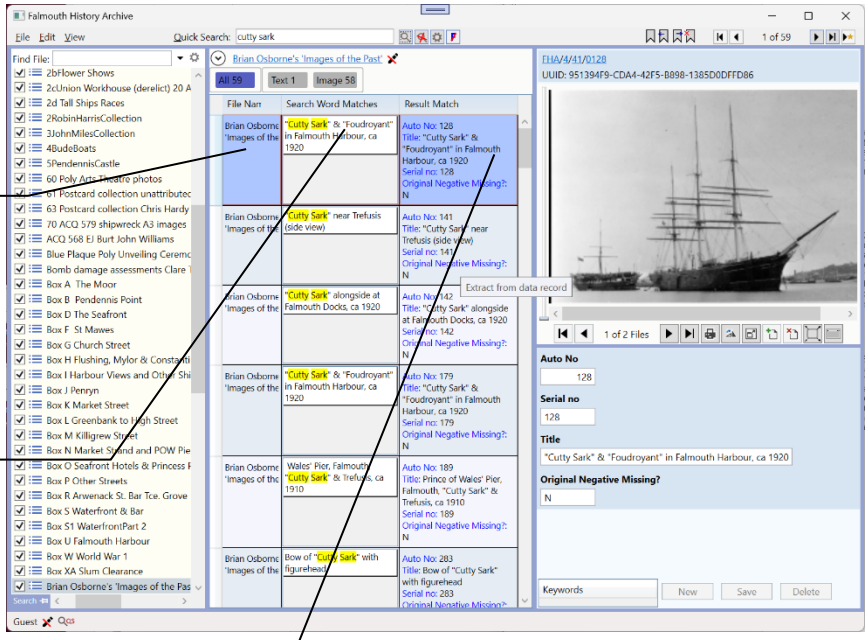
If you have understood the message and do not want to be reminded of this again, check the box *Do Not Show This Message Again*. Then click OK.



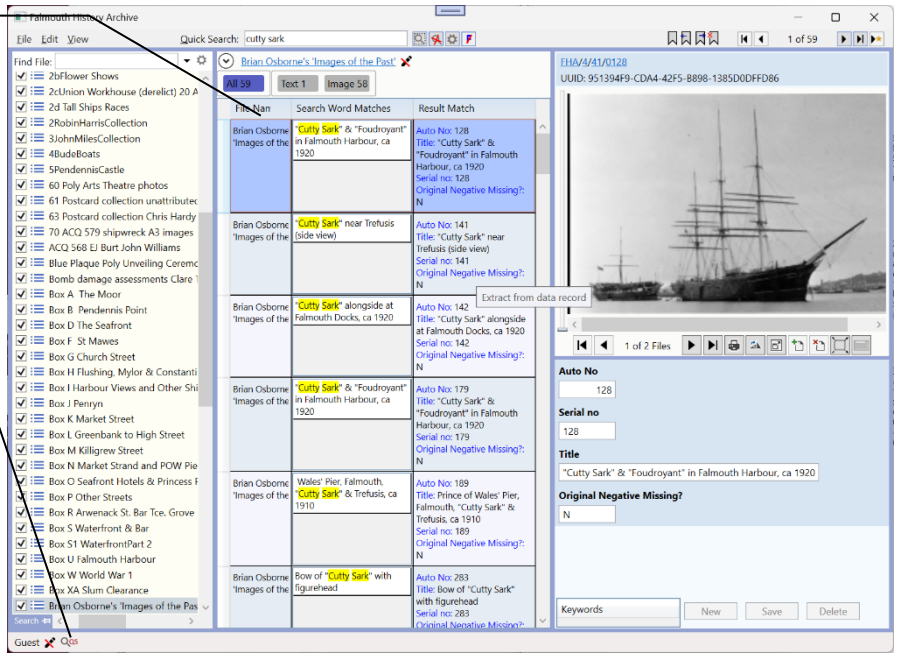
Results are returned in a grid with 3 columns. The first column shows the name of the **File** containing the returned record.

The 2nd column contains a grid that lists matches to the search criteria within the record, along with a little text either side of the word match. The matches are highlighted in yellow

The 3rd column is a summary of the data held in the matched record. The format for the data displayed is the display name for the field (in blue), followed by a colon followed by the beginning of the data stored in that field (e.g. *FIELDNAME: This is some data*)



Results are returned in natural order but can be re-ordered by clicking on the column headers, or from sort options on the right-click context menu or in the Quick search options dialog.
See *Sorting and Grouping,* below.

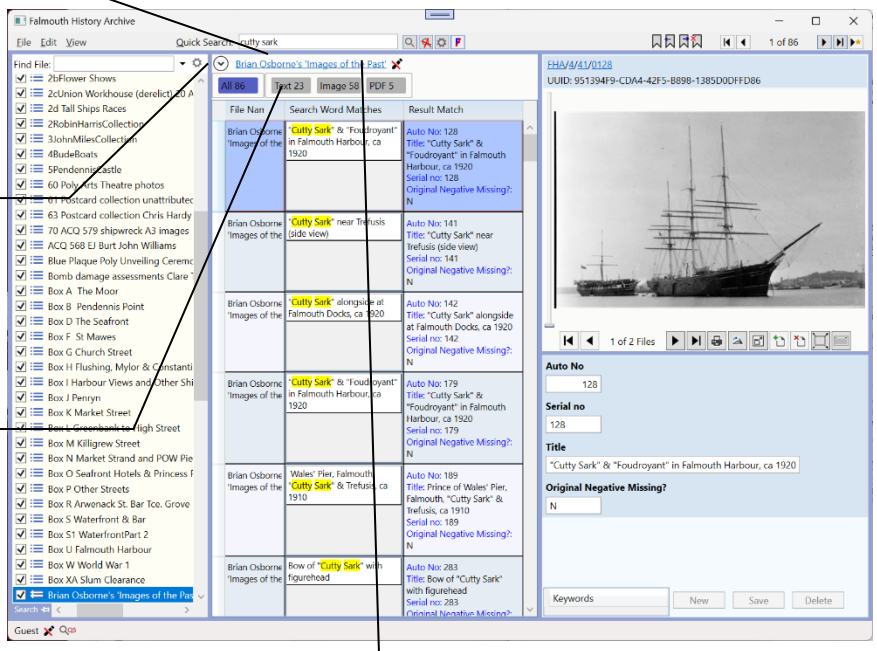


This icon indicates the current data mode is Quick Search

The blue hyperlink above the result grid is the file name for the current record. Click the hyperlink to browse the file from the record selected in the QS results

To the left of the hyperlink is a small down arrow. Click this to display current file information.

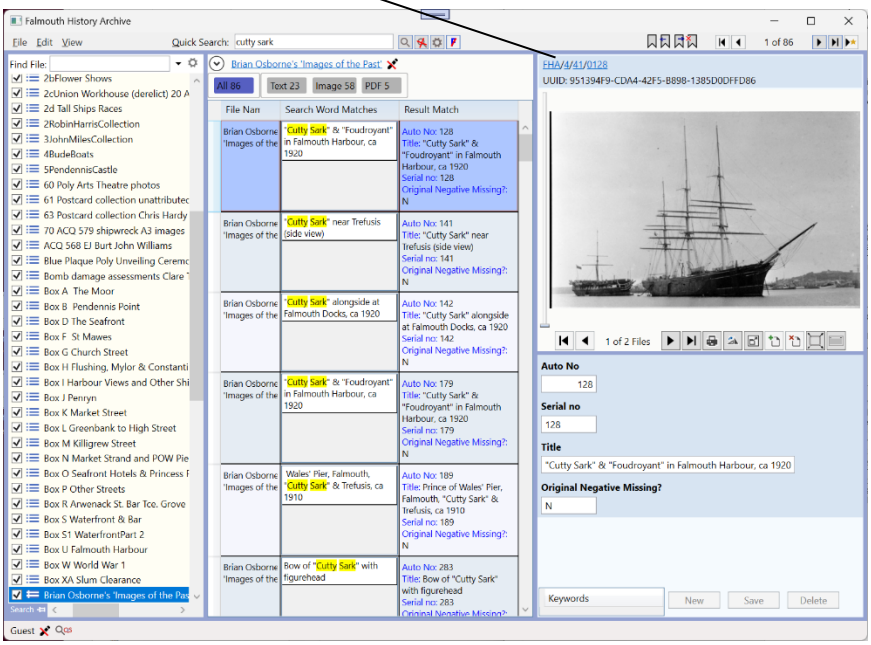
Immediately above the result grid are the filter buttons. These can be toggled to filter the results. In this example there are 23 records containing text only, 58 records containing images and 5 records containing PDFs. The **All** button clears your filter selection. The other filters are additive.



To the right of the hyperlink is an icon that indicates the current user access rights to the selected file.

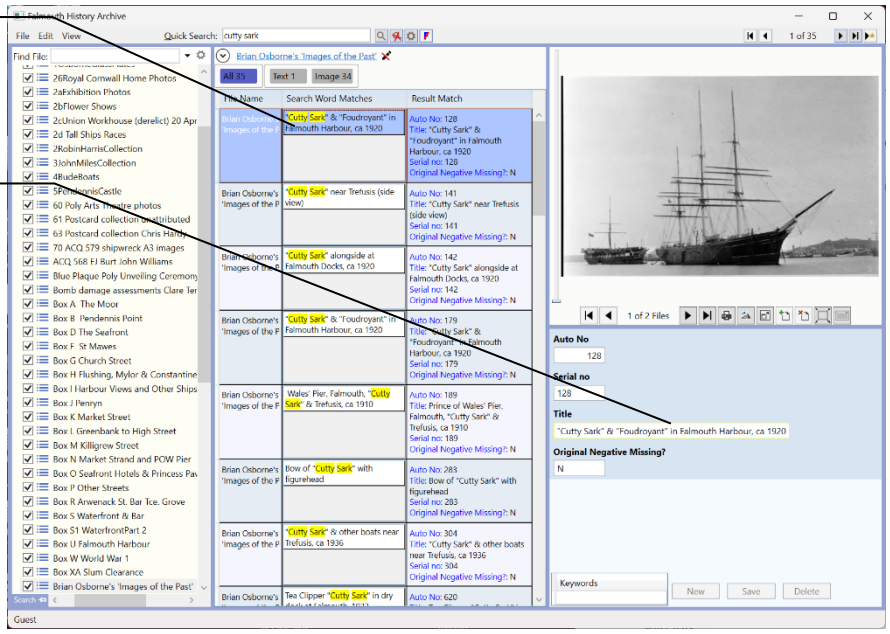
The results of the Quick Search are read-only: if you wish to edit the record, either click the file-name hyperlink or the current record (right-most) component of the Hierarchy Reference Number. This will take you to the record in the Browse Grid.

The **Hierarchy Reference Number** hyperlink is shown here. Its visibility can be turned on and off by an administrator setting.



Clicking a match grid cell…

…will place a border around the field containing the match data

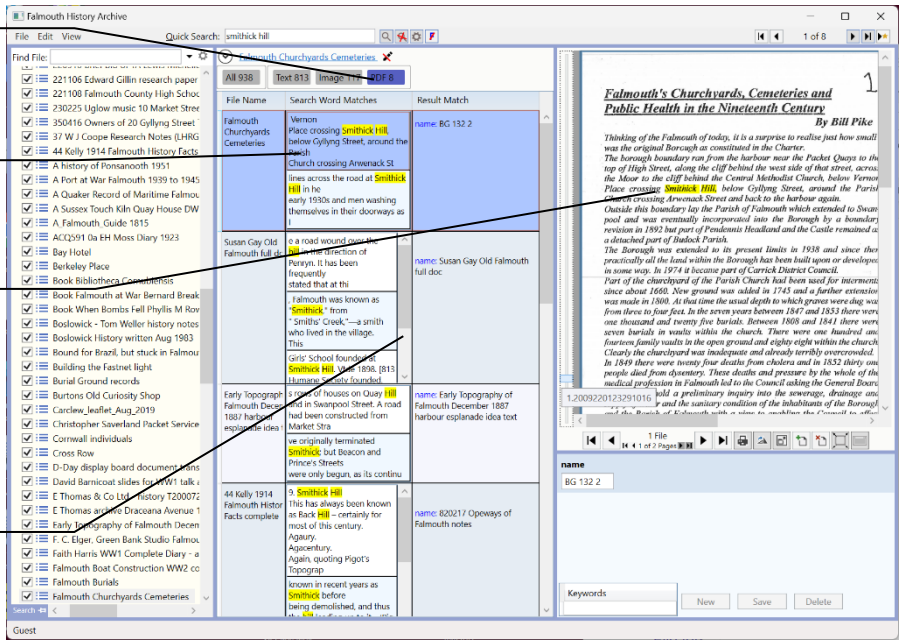


In this example results have been filtered to show only records containing PDFs

Clicking a match grid cell for a PDF result…

…will navigate to the appropriate PDF page and highlight the match. Note: if the PDF pane is smaller than the PDF page, the highlighted match may be out of sight. Enlarge the pane.

Each of these result records contains many matches and the inner match grid can be scrolled.



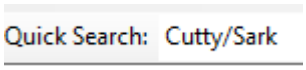
Quick search criteria

The quick search is designed to fulfil most search requirements. It is a full text search of data records, PDF text and Audio Visual Transcripts. It is a global[2] word and date search, not an individual field search. It requires the search terms to be entered with a defined syntax. Please note: I have capitalised most of the search criteria for clarity but it is not necessary to enter search terms in uppercase. All searches are case insensitive. The search does not accept wild cards.

It is not necessary to know the structure of the tables and you cannot specify individual fields in the search. The Quick Search searches data held in any text or date field. Numeric and geolocation data cannot be searched using the Quick Search. Dates can be searched by whole years, years and months or specific dates (dd/MM/YYYY).

The search syntax is reminiscent of an algebraic formula. In its simplest form, however, the search criteria can simply be a few words as in: "*CUTTY SARK".* The search interprets this as "show me records containing the words *CUTTY* AND *SARK"* and looks for records that contain both words, *CUTTY* and *SARK.* "*CUTTY AND SARK"* is an 'AND' criterion. The default is for the search to interpret a string of words as *AND* criteria. Because this is the default you can omit typing 'AND' between words in simple AND searches. A '+' sign can be used in place of typing *AND*, as in "*CUTTY + SARK"*.

The search can also use *OR* criteria as in "*CUTTY OR SARK*". A '/' (slash) can be used instead of typing *OR*. "*CUTTY/SARK*" will return all records containing either the word *Cutty* or the word *Sark.*

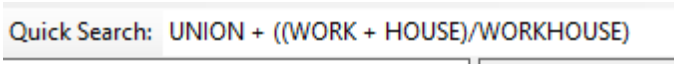


[2] Global in this context means the whole database archive, or a selected sub-set

The search is an exact word match. To search for alternative word forms, spellings and plurals, you can combine *AND* with *OR* criteria in an algebraic like formula, as follows. In this next example we want to find records relating to the *workhouse*. *Workhouse* can be written as the single word *WORKHOUSE* or as two words: *WORK HOUSE.* The search term required to return both forms is (omitting the quotes):
"(WORK + HOUSE)/WORKHOUSE"

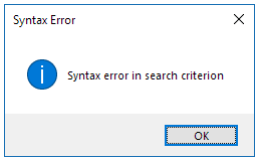
Note the brackets that create 2 levels in the criteria; we are requesting either *WORKHOUSE* OR (*WORK + HOUSE)* together.

This can be extended: say we want to search for UNION WORKHOUSE, or UNION WORK HOUSE, we can bracket around the above criteria, as in *((WORK + HOUSE)/WORKHOUSE),* and then add UNION, as follows (again omitting the quotes) to create a 3 level criteria:
"*UNION + ((WORK + HOUSE)/WORKHOUSE)"*



When entering complex search terms it is mandatory to include **+** and **/**.

You will also see the syntax message, below, if you incorrectly bracket search terms or if you fail to alternate AND with OR criteria levels: *"(WORK + HOUSE)+WORKHOUSE"* is meaningless and will generate a syntax error.



Beware of plurals and alternative spellings: 'GARDEN' and 'GARDENS', for example, are different words and you might wish to look for both by using an OR criterion e.g. '*GARDEN/GARDENS*' or by using a Fuzzy Search.

A further search example might be: "*(JOHN/JON/JONATHAN)+(SMITH/SMYTH/SMYTHE)"* to look for people matching all name forms for John Smith. There is no specific limit to the number of nested criteria but, beyond a certain point, it becomes very difficult to understand your own search request.

Searching for dates within the Quick Search

To specify that you are searching for a date, enclose your date criteria between exclamation marks, as in: *!28/02/2018!*. The format for date criteria is relatively flexible; the following criteria is equally acceptable:*!28 FEB 2018!*. Day criteria must be first, month next and finally year. Always use a 4 digit year. These examples search for a specific date, but criteria can be less specific. It is possible to search for any date within February 2018 as follows: *!FEB 2018!*. Any date in 2018 can be searched with: *!2018!*.

Dates entered as above can be combined with other criteria. For example "*UNION+!JAN 1886! + ((WORK + HOUSE)/WORKHOUSE)*" (omitting the quotes) will return 1 result for someone with a date of birth of 6 Jan 1886 at the union workhouse.

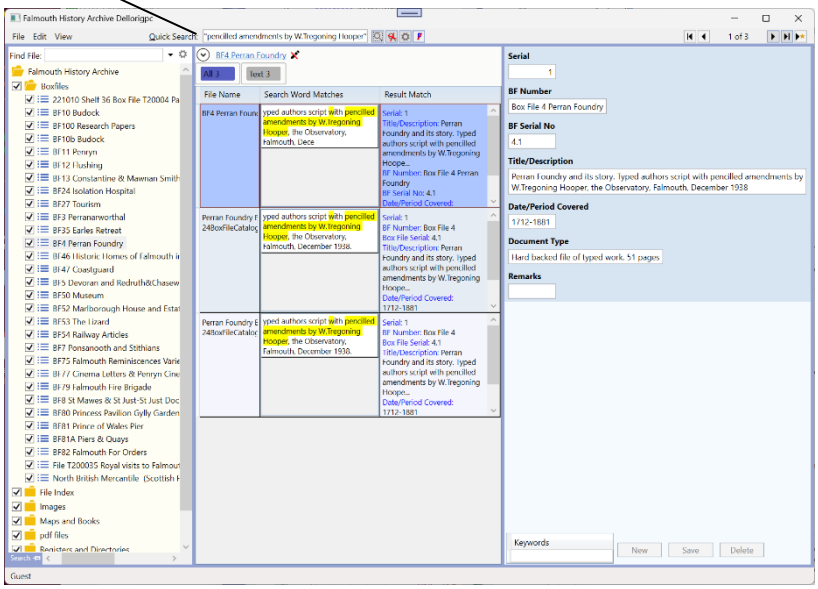
Searching for Keywords within the Quick Search

File-based keywords should be preceded by the '@' symbol. This tells the search that the word is a File keyword. Record-based keywords should be preceded by the '#' symbol. An example could be "*@BOXFILES/@BOXFILE*" to search for all records within files that have the keyword *BOXFILES* or the keyword *BOXFILE*. Keyword criteria can be combined with other criteria in the algebraic Quick Search as in: "*(@BOXFILES/@BOXFILE) +PERRANARWORTHAL*" to find all records having *PERRANARWORTHAL* in files also having the keywords either *BOXFILES* or *BOXFILE*.

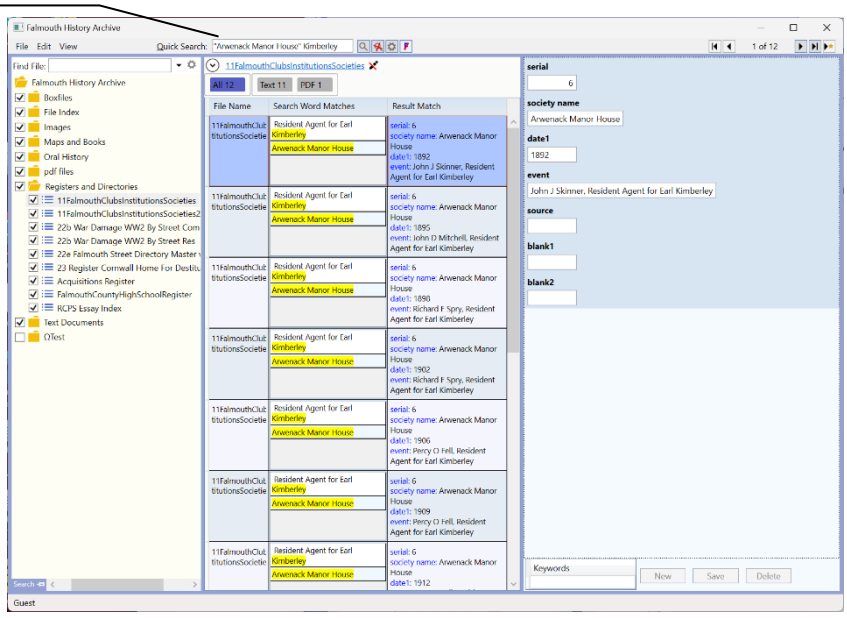
Quoted Text Search

You can search for an exact match to a phrase by enclosing it in double quotes.

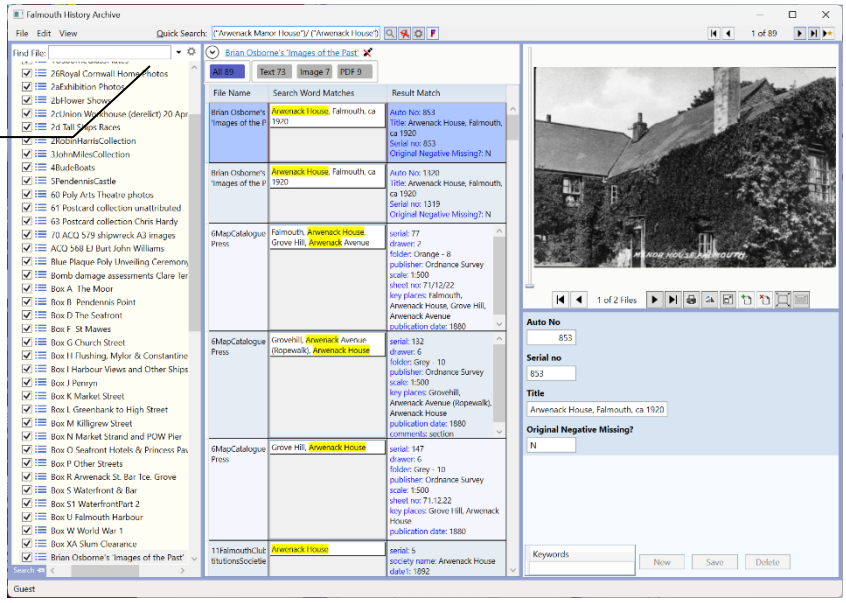
Surround your search text with double quotes to search for an exact match to a phrase.

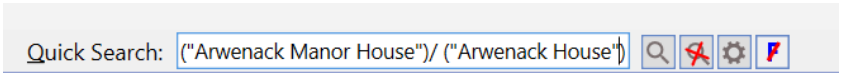


Quoted criteria, words or phrases, can be mixed with unquoted words within an algebraic search, eg *"Arwenack Manor House" Kimberley*
This search will return records containing the phrase *"Arwenack Manor House"* along with the word Kimberley within the same record



If you combine 2 quoted text criteria in an OR search, eg *"Arwenack Manor House"/ "Arwenack House"*, you must enclose each OR criterion in parentheses with the slash representing OR between the brackets, as shown below.

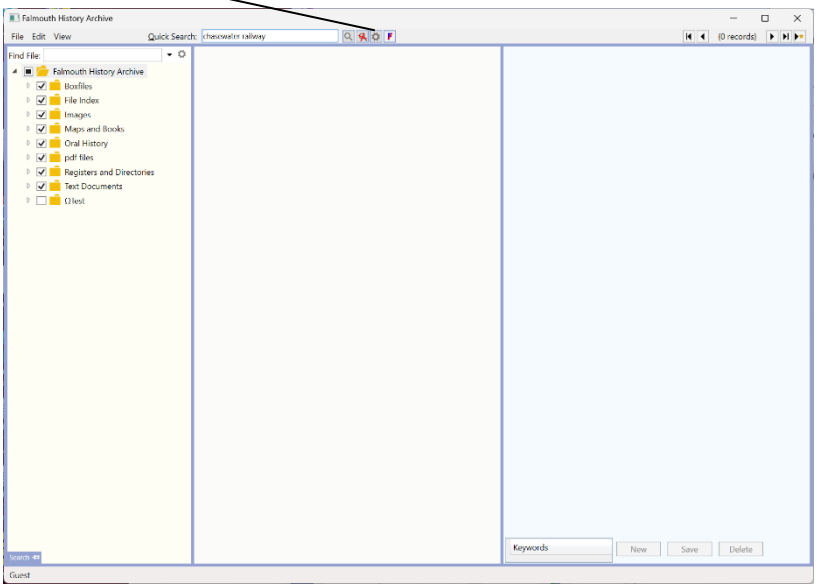




Fuzzy matching

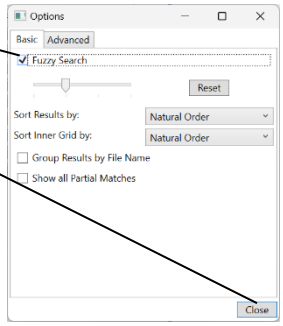
To turn on Fuzzy matching do one of the following:

Open the Search Options (cogwheel button)



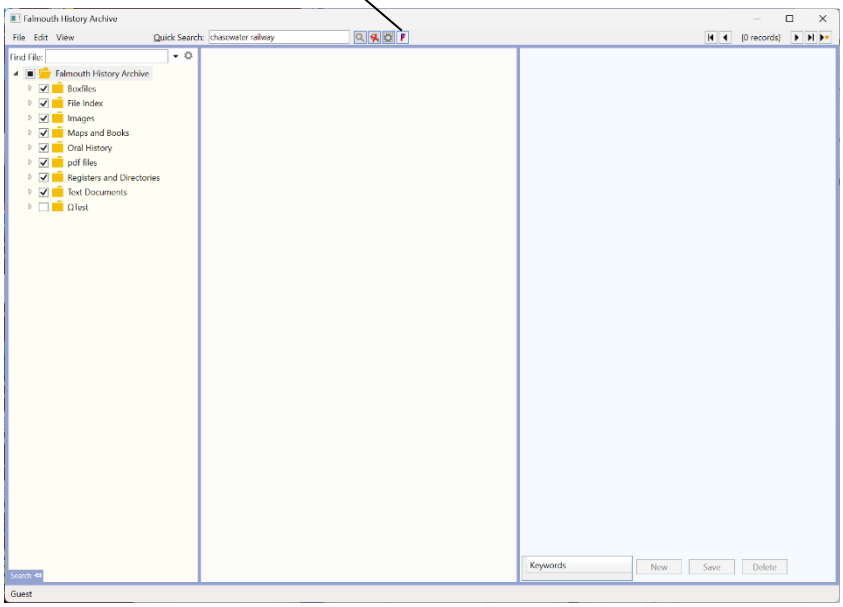
Check **Fuzzy Search** and…

…and close the dialog (You can click or hit the return key)

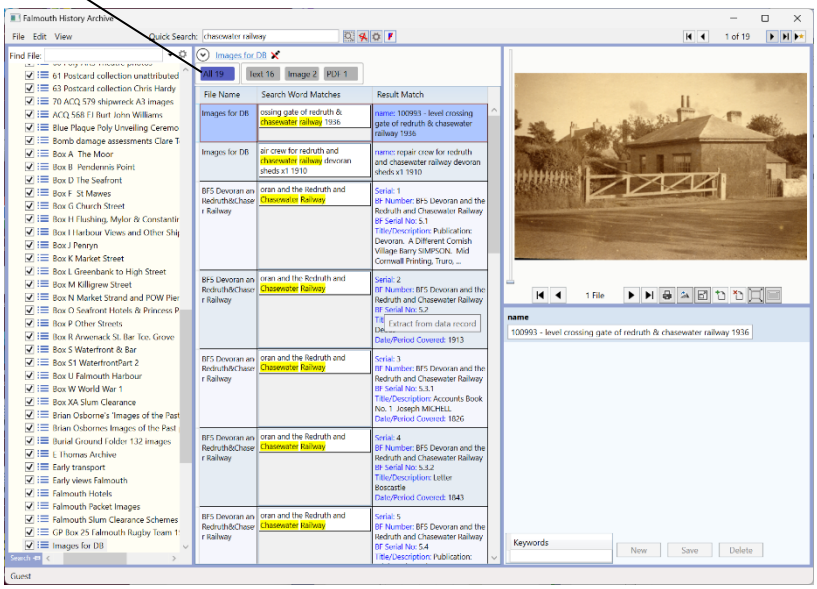


OR

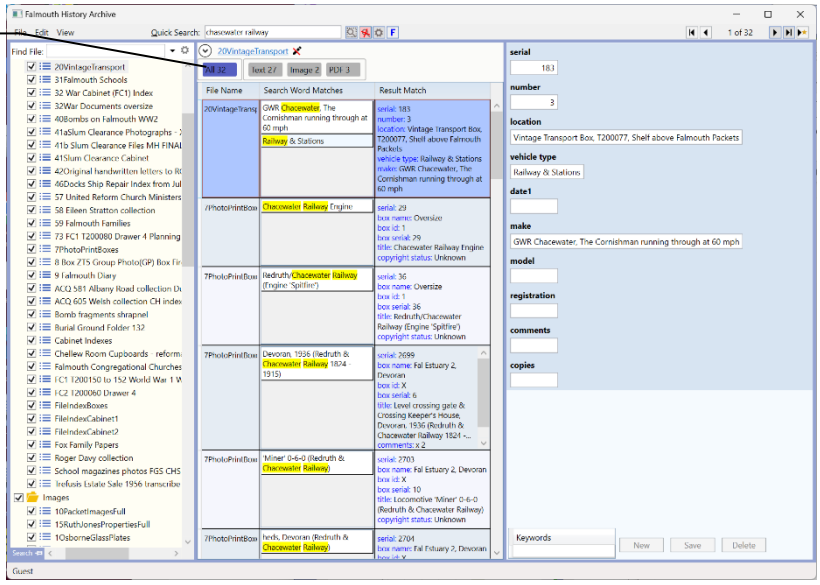
Click on the Fuzzy toggle button. This will toggle the Fuzzy Search on and off



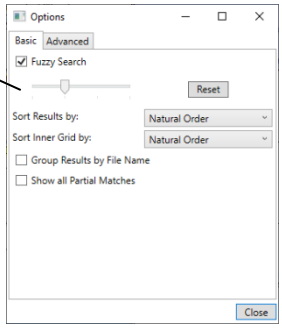
These are the results of a non-fuzzy search for *Chasewater railway* 19 records are returned



Turn on Fuzzy Searching and repeat the search: 32 results are returned including *Chasewater* and *Chacewater*



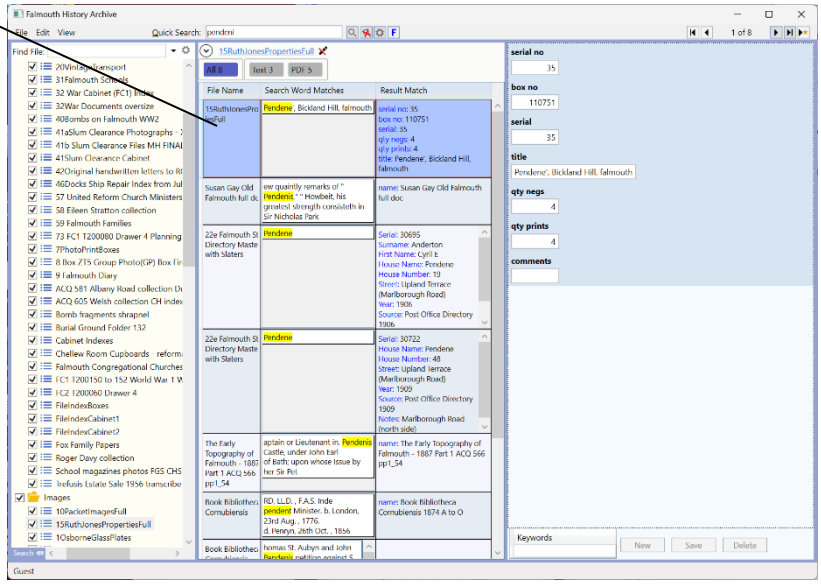
The Options dialog has a slider that changes the 'looseness' of the match



The numerical value can be between 1 and 4 and represents the number of character substitutions permitted. Searches become slower with longer search texts and 'fuzzier' searches.

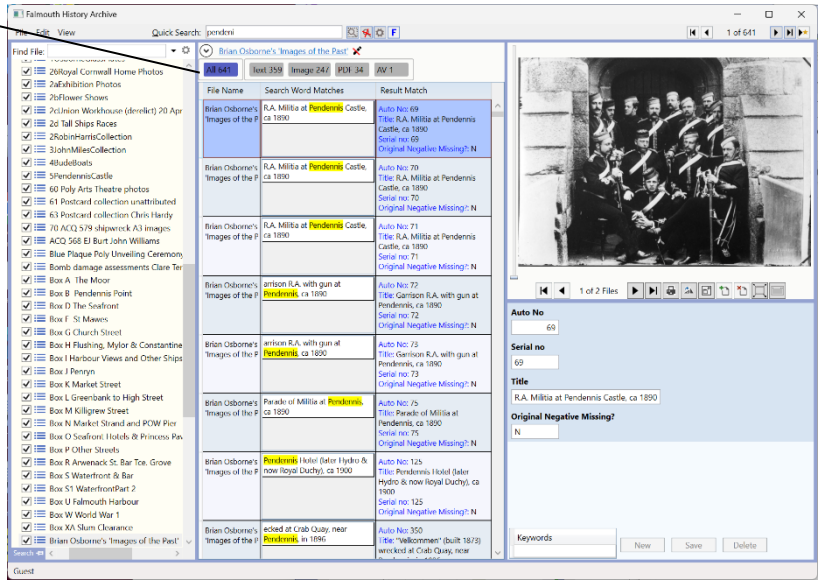
You can specify that the search excludes specific words from the fuzzy search by enclosing then in double quotes: *(("JOHN")/("JONATHAN")) +SMITH*[3] with the slider moved to 3 will return only exact matches to JOHN and JONATHAN along with a variety of SMITH spellings

The slider is set to 1
The search is '*pendeni*' and the results brought back include '*Pendene*', '*Pendenis*' and '*Penden*'



[3] Note the brackets around *("JOHN")* and *("JONATHAN")* and then around the whole OR criteria: *(("JOHN")/("JONATHAN"))*

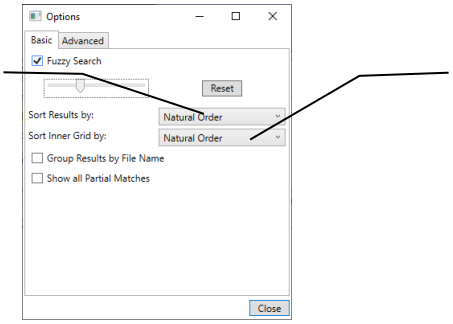
With the slider set to 2 and the results include '*Pendennis*'



Sorting and Grouping

Results are returned in **Natural Order** by default. You can also order results by:
Average Matches
Total Match Score
File Name ASC
File Name DESC

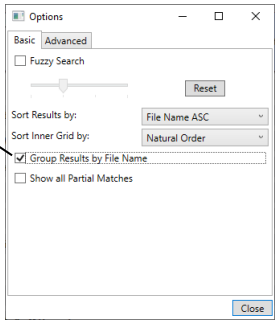
Average Matches is the average score of each result within a record. This option tends to favour best matches for shorter text records.
Total Match Score is the total score of a record and this option tends to favour records with many matches, e.g. PDF's.
The file name options are self-explanatory. Once set, sort options continue to apply to subsequent searches.



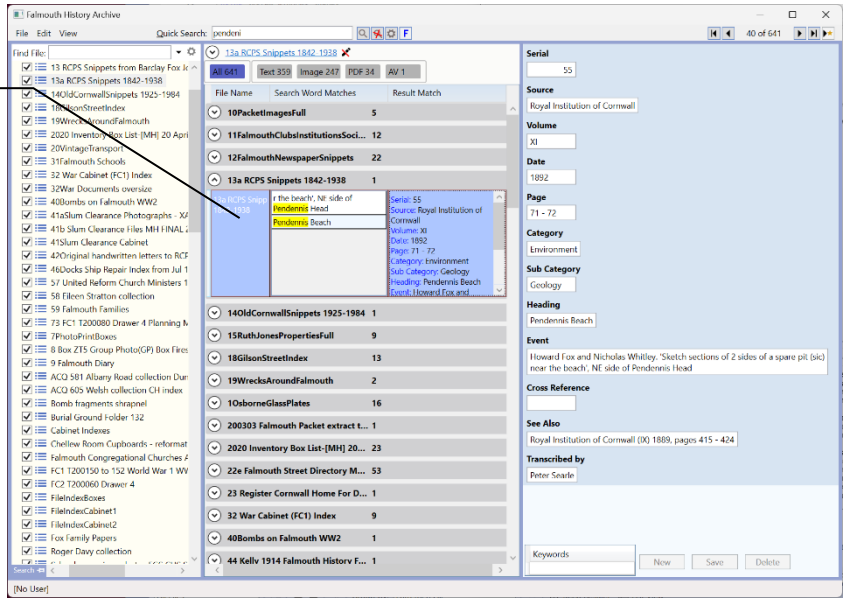
Inner Grid sort options are **Natural Order**, which orders items as they appear in the record, or **Best Matches** that orders inner grid items by match score. This will make no difference if you use a single word criterion

Grouping by File Name

Check the **Group Results by File Name** check box to group results. The option is also available on the right-click context menu for the search grid

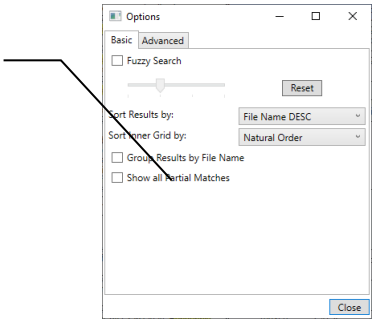


Grouping collapses the results into grey headed groups. To the left of each header is an expander glyph to expand the group. When you group records, the group containing the current record is expanded and the rest are collapsed. File ordering, ascending or descending, also applies to groups.
Grouping slows performance and can only applied after running a search and so will not apply to subsequent searches.



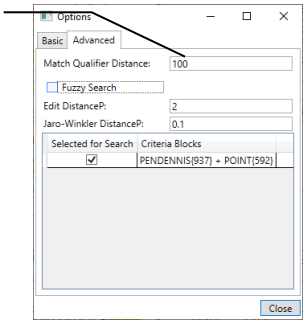
Advanced search options

Show all Partial Matches returns all partial matches to compound matches – that is searches having more than one ***and*** criteria.
By default, poor partial matches are ignored to show the best matches that are likely to be the most relevant. If you are concerned that the search is missing something you can check this option and search again. The search must be re-run to see the difference. It is recommended to keep this option off.



Advanced Tab

Match Qualifier Distance is the number of words between match word candidates to qualify for a match. You can set it between 20 and 1000.



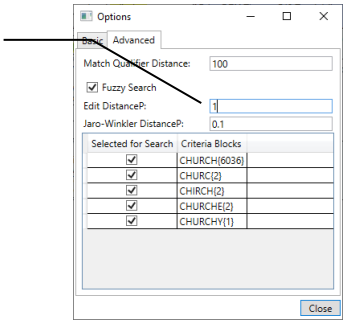
If you were searching for *Market Street*, say, with a qualifier distance of 100, the word *Market* would form a match with next or previous word *Street* if they were up to 100 words apart.

The value to choose depends on the context of your search. For *Market Street* you would likely opt for a small distance or perhaps a quoted text search. If you were looking for, say, Beacon Street in

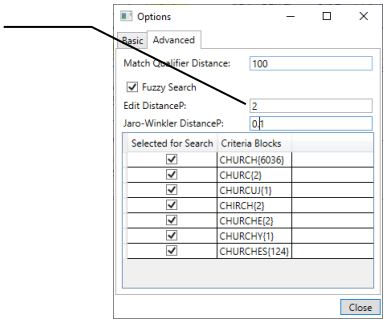
relation to slum clearance, you might opt for a greater **Match Qualifier Distance** to reflect a looser word association. Your search might look like: "*slum +(clearance/clearances) + beacon + street",* omitting the quotes

Changes to **Match Qualifier Distance** apply to subsequent searches but are not reflected in any current results.

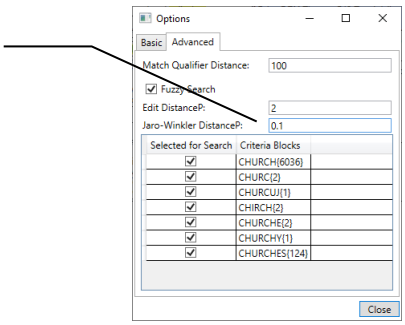
Edit Distance is the number of character changes allowed by the fuzzy search. The grid below the text boxes lists the words that will be searched for your chosen parameters. This example shows the result for a criterion of ***church***. Note that ***churches*** is not included in the list. ***Churches*** is 2 character edits away from ***church***.
Edit Distance can be any value between 1 and 4.
Changes apply to subsequent searches.



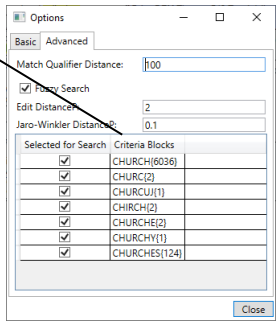
Changing the Edit Distance to 2 adds ***churches*** to the list.



Jaro-winkler Distance is a measure of word similarity. The default is 0.1. Valid values are 0 to 1 with 0 being most similar.
When you move the slider on the Basic Tab, the value of the **Jaro-winkler Distance** is increased with the increased **Edit Distance**. You can use the Advanced Tab to set the value manually and independently of Edit Distance.
Changes apply to subsequent searches.



The grid below the parameters displays the words that the fuzzy search will look up. You can create a bespoke search by removing the checks in the **Selected for Search** column. Each word is followed by (within braces) the word count in the database records.
The grid has a context menu that allows you to check or un-check all the boxes.

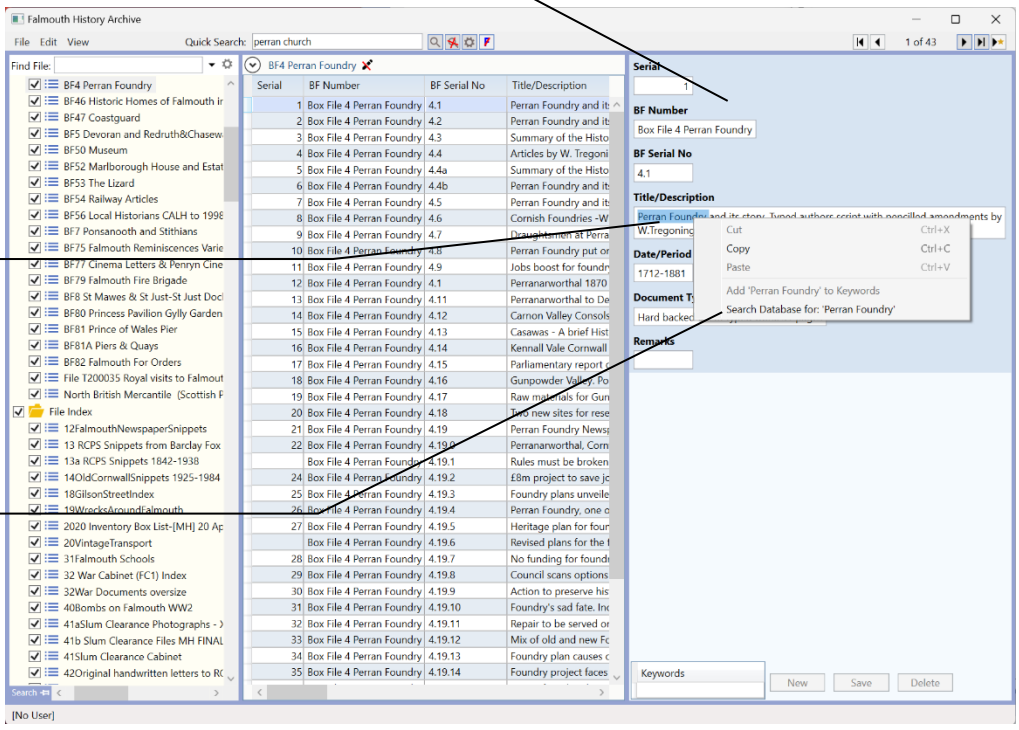


Shortcut Quick Searching

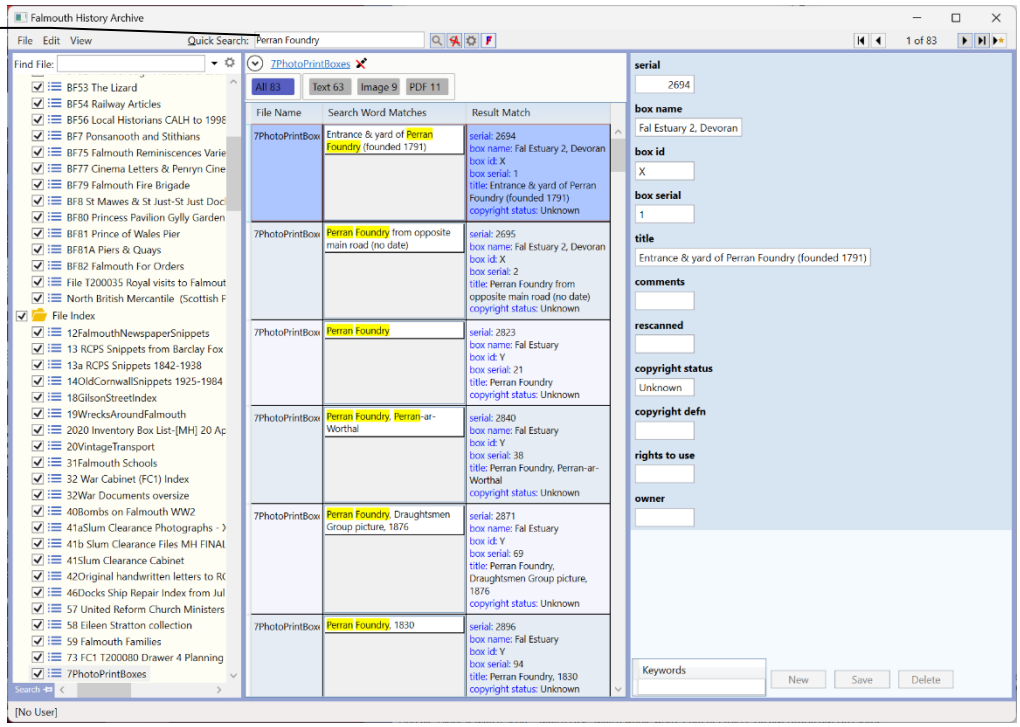
If you are browsing a record and wish to find records that relate to certain text within a field...

…select and right click the text to display the context menu…

…choose **Search Database For:** from the context menu…



...the search terms will be added to the **Quick Search** box and the quick search will run to display the results



Detailed searches

Introduction to Schemas and Data Types

All File data in ArchiveDb is stored in tables. Tables consist of rows and columns. Columns can also be referred to as ***Fields***. Each column (field) has a name and caption and each column stores a particular type of data. ArchiveDb currently has 11 different data types to choose from. Examples of data types are: text, dates, numbers, binary and location data. The 11 different data types will be discussed in detail, later.

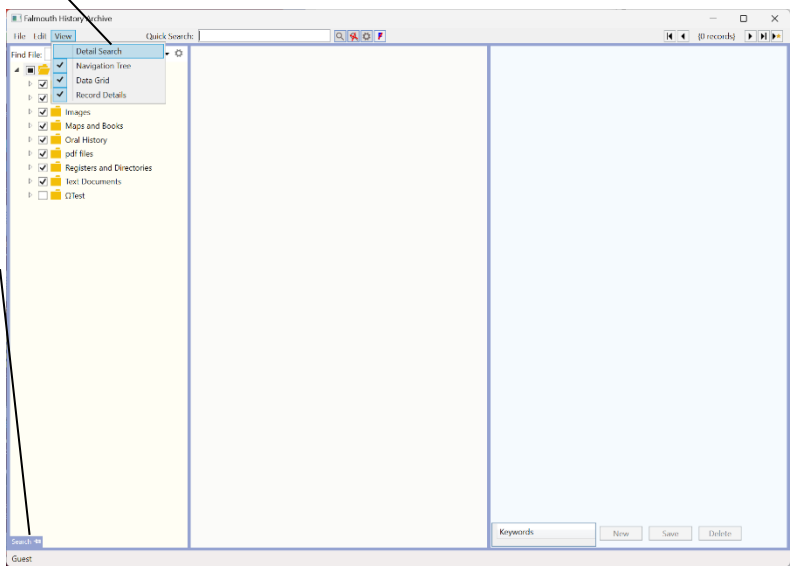
The column specification for each table is called the ***Schema.*** Every File must have a Schema but a Schema can be shared between several Files, thus several Files may have identical table structures (same column names and data types etc).

The detailed searches allow you to search individual Files, and multiple Files that share the same schema, by their individual, named, fields.

The detailed searches also allow you to search a group of Files that have more than one schema, this is a multi-schema search. You cannot always search individual fields if you are searching across multiple schemas; you may only be able to search by individual data type.

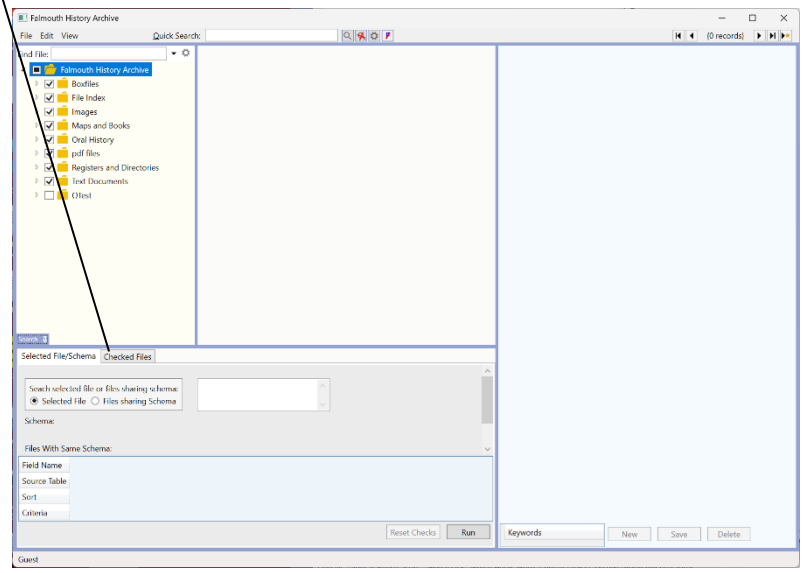
Searching single files

The detailed search is hidden by default. To view the detailed search pane either check the **Detailed Search** in the **View Menu** or…



…click on the **Search** toggle button at the bottom of the navigation tree.

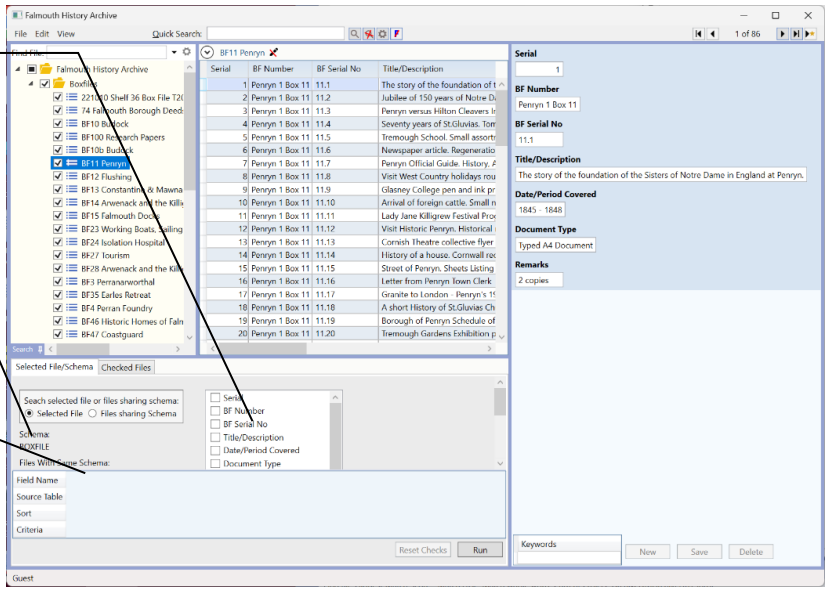
This is the search pane. There are 2 tabs: **Selected File/Schema** and **Checked Files**.
In this screen print no file has been selected



Having selected the file its fields are listed as check boxes in this list.

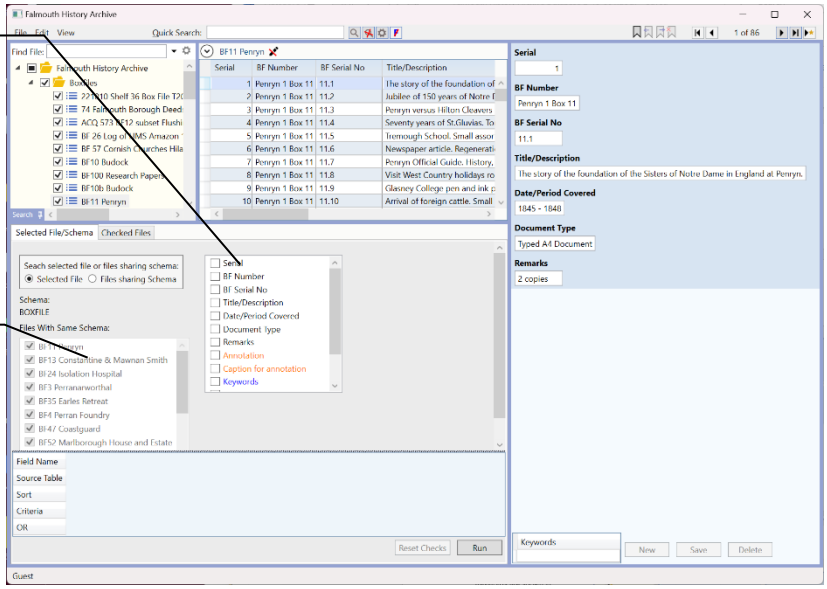
The Schema name is shown here.

You can drag this divider to enlarge the panels, as required.



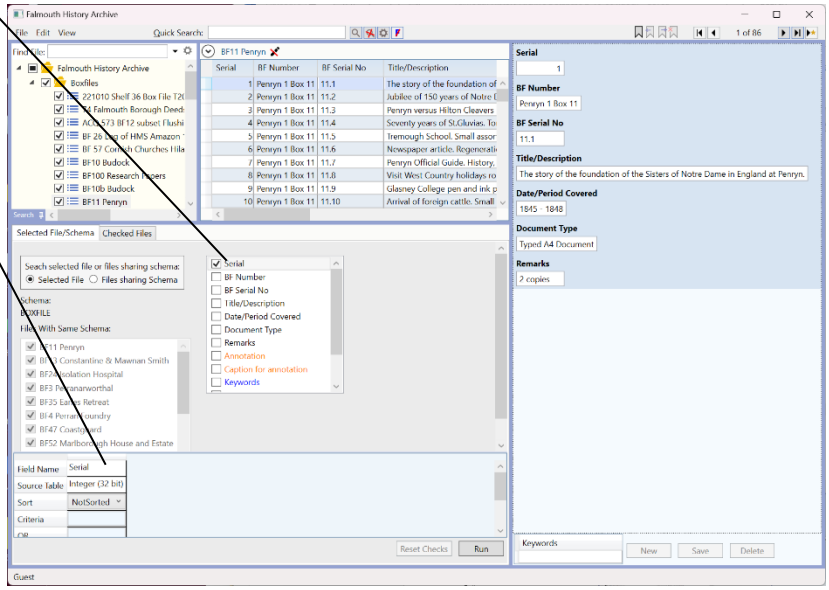
User defined fields are shown in black. Annotation fields are shown in orange, Keywords in blue and the system fields (that belong to every file) in brown. To apply criteria to a field you check its check box

This is a list of files that share the same schema. It is greyed out (disabled) when searching the selected file



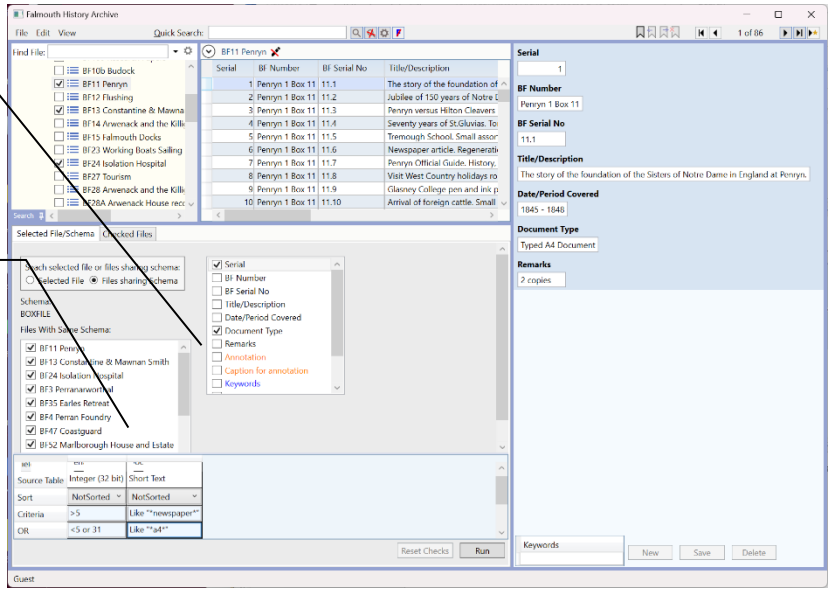
Checking *Serial*…

…adds the field to the query grid.



I have also clicked *Document Type* to apply criteria to that field.

Both checked fields appear in the query grid. Fields are added in the order they are clicked. This is relevant when you apply a sort to sort by more than 1 field as the fields are sorted preferentially in the order in the query grid.



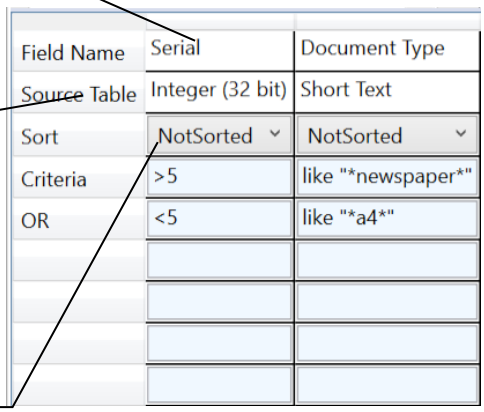
The results of single file searches can be edited if you are logged in and have edit rights for the file.

The query grid

The query grid is reminiscent of the Microsoft Access query grid, although not as fully functional: specifically, it does not support drag and drop operations and does not support table joins and many query expressions are unsupported. It also differs in a number of other respects.

Field Name refers to the field in the user defined table schema. In this example that table is '*BF11 Penryn*' and the specific schema is '*Boxfile*'

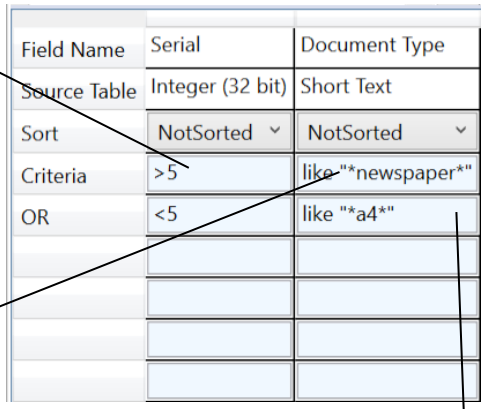
Unlike the Access query grid, **Source Table**, however, refers to the underlying ArchiveDb table that stores data of the relevant field type. **Source Table** is, in effect, the data type.



Sort is the sort order you specify for the field in the resulting query. A sort can be specified for more than one column and the query will be sorted using the sort orders read from left to right.

Query grid criteria

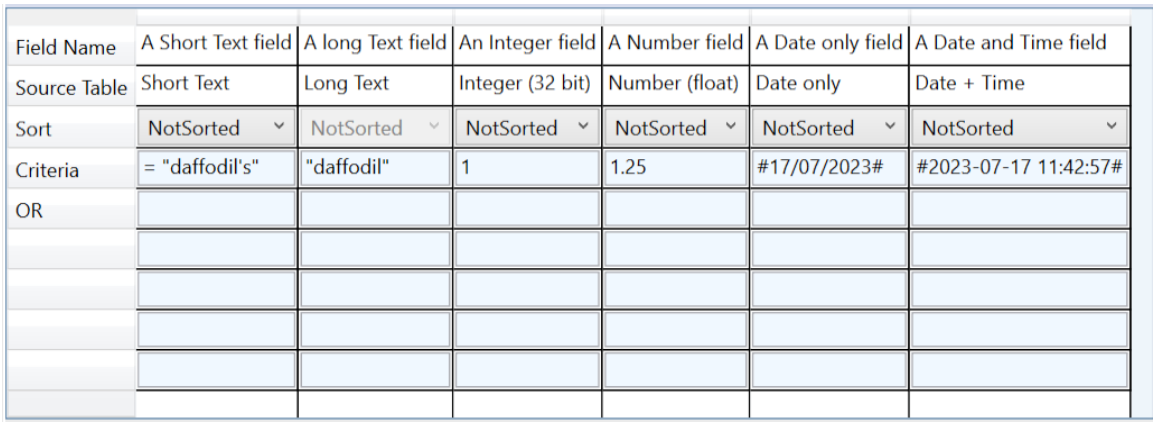
You add your criteria to the criteria cell of the query grid.



Criteria added along a row are **Add** criteria. In this example this row specifies rows having a *Serial* number >5 AND a *Document Type* that contains the text *newspaper*

The next row is headered **OR**. Criteria added to the this and other rows are alternative criteria combined with OR.
In this example the query will return rows having a *Serial* number greater than 5 AND also contain '*newspaper'* in the *Document Type*, OR rows that have a *Serial* number less than 5 AND also contain the text '*a4'* in the *Document Type* field.

Query grid criteria for specific datatypes



Field Name	A Short Text field	A long Text field	An Integer field	A Number field	A Date only field	A Date and Time field
Source Table	Short Text	Long Text	Integer (32 bit)	Number (float)	Date only	Date + Time
Sort	NotSorted	NotSorted	NotSorted	NotSorted	NotSorted	NotSorted
Criteria	= "daffodil's"	"daffodil"	1	1.25	#17/07/2023#	#2023-07-17 11:42:57#
OR						

The example query grid, above, illustrates the syntax for applying equals (=) criteria to the basic, (non-compound) datatypes. Location and Decimal are compound datatypes.

Operators define how you wish to make a comparison: equals, greater than, less than, between etc.

The default operator is equals (=) and can, optionally, be omitted. Compare the first two sets of criteria in '*A Short Text field*' and '*A Long Text field*' for "daffodil's" and "daffodil".

Short Text and **Long Text**: text criteria must be surrounded by quotes. Quotes can be double or single. If you are searching for text that includes single quotes (i.e. your criterion has single quotes within) you must surround text with double quotes. If your criterion contains a double quote you must surround it with single quotes. It is therefore not currently possible to use a criterion that contains both double and single quotes.

Integers and **Numbers**: numeric values should be entered as-is, without any quotes.

Date only and **Date + Time**: as with the Quick Search, date criteria must be surrounded by **!** (exclamation marks). For compatibility with the Microsoft Access query grid, you may also use **#** (hash symbol). **Date only** criteria can be formatted more flexibly than **Date + Time**. You can use slash **/**, dash **–** and spaces to separate the date parts and the parts can be formatted as either yyyy/mm/dd or dd/mm/yyyy. Thus *!2022 11 21!* is equivalent to #21-11-2022#. **Date + Time** however has very specific format requirements for criteria: date and time must be entered as yyyy/mm/dd hh:nn:ss. This can be surrounded by either **#** or **!** symbols. **nn** is the shorthand for minutes in 2 digits.

Comparison operators

The following operators are valid:

= *equals*
> *greater than*
< *less than*
>= *greater than or equal to*
<= *less than or equal to*

Between *val1* And *val2* *Between 2 values*
Is Null *field contains no value*
Is Not Null *field contains a value*
Not *field does not contain specified value. Equivalent to* **<>**
<> *field does not contain specified value*
val1* OR *val2 *field contains val1 or val2 (etc)*
val1* AND *val2 *field contains val1 AND val2 (etc)*
Like *specific to text and wildcards. See below.*
Not Like *specific to text and wildcards. See below.*
Microsoft Access also has the **IN** and **NOT IN** operators. These operators are not supported by ArchiveDb but the functionality can be replicated as follows: 'IN (1,2,3)' can be replaced by '1 OR 2 OR 3'. 'NOT IN (1, 2, 3)' can be replaced by the rather counterintuitive 'NOT 1 AND NOT 2 AND NOT 3'

Wildcards

Wildcards are used in short and long text field searches. Wildcards represent unspecified characters. ArchiveDb accepts 2 wild cards: **?** and *****, the question mark and the asterisk.

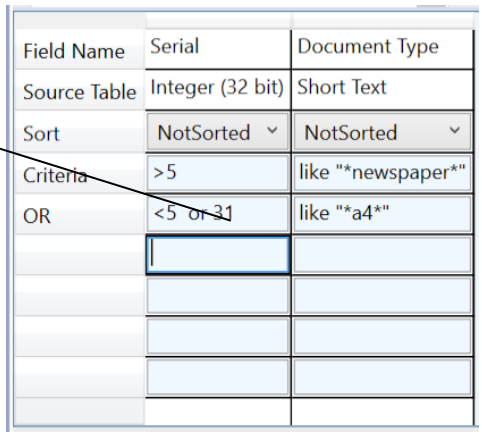
The question mark represents exactly 1 unspecified character and the asterisk zero or more unspecified characters.

Wildcards are used with the **Like** operator. An example criterion would be '*Like "*newspaper**"*' and this would find records that contained the text '*newspaper*' somewhere within the field. The asterisk represents zero or more characters so this criterion would find records with '*newspaper*' anywhere in the field, including the beginning and end and including being the entire field content. It would find '*newspapers*' as well as '*newspaper*'.

Combining criteria

Complex criteria can be built by combining criteria using the **AND** and **OR** operators within a query grid cell.

Criteria can be combined within a query grid cell. In this example the same records will be returned as in the earlier example but with the addition of the record having a *Serial* value of 31, provided that its *Document Type* contains the text 'a4'



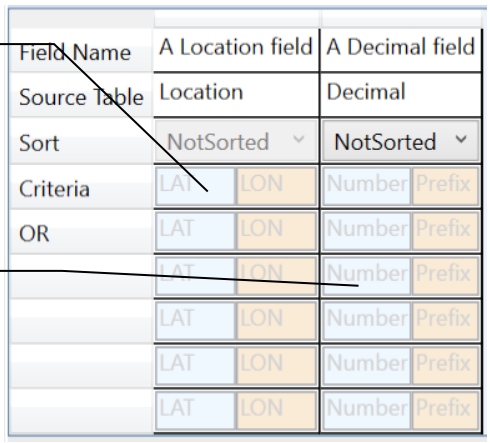
Criteria for complex datatypes: Location and Decimal

The **Location** datatype consists of Latitude and Longitude data, the **Decimal** datatype consists of a prefix and a number. See datatypes, below.

The query grid column criteria cells for these 2 datatypes are divided in two.

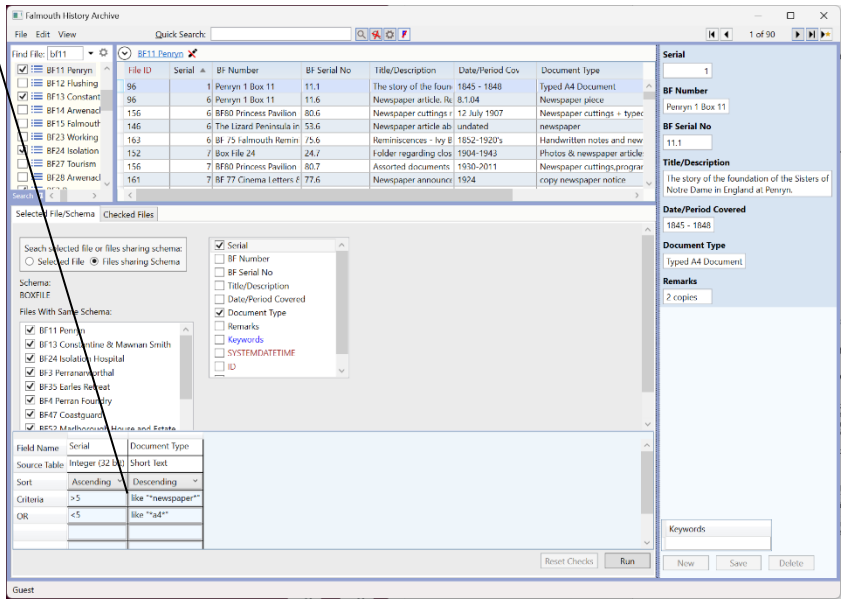
Criteria for **Location** data can be applied to both latitude and longitude. These are **Number** subfields. The query cells have watermarks to indicate the sub-field for the criteria.
At the time of writing **Location** data cannot be sorted.

Criteria for **Decimal** data can be applied to both number and prefix. Note that the query criteria have the **Number** criteria placed ahead of the **Prefix** criteria. At the time of writing the sort is applied only to the numeric value of the field, disregarding the prefix.



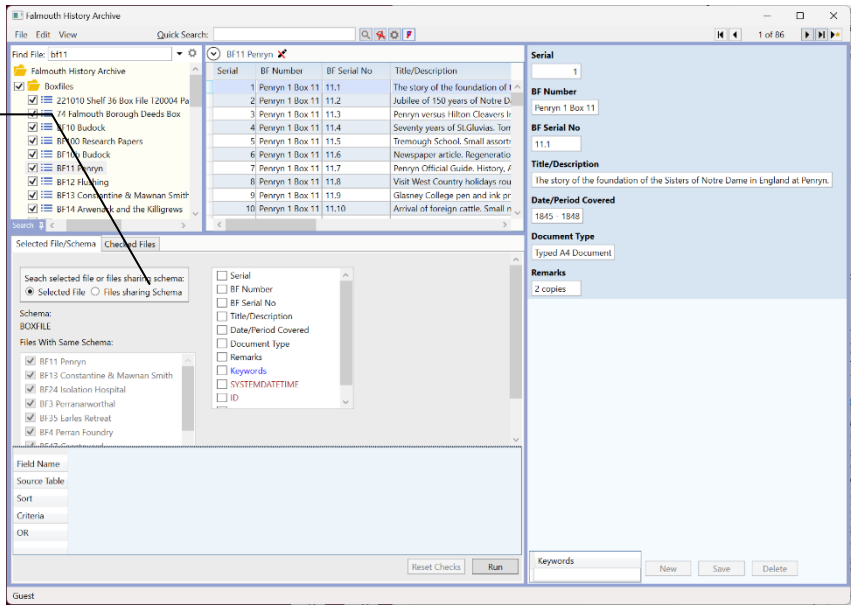
Sc

Results can be sorted with the query grid. In this example the results will be sorted by *Serial* in ascending order and then by *Document Type* in descending order.



Searching multiple files sharing a single schema

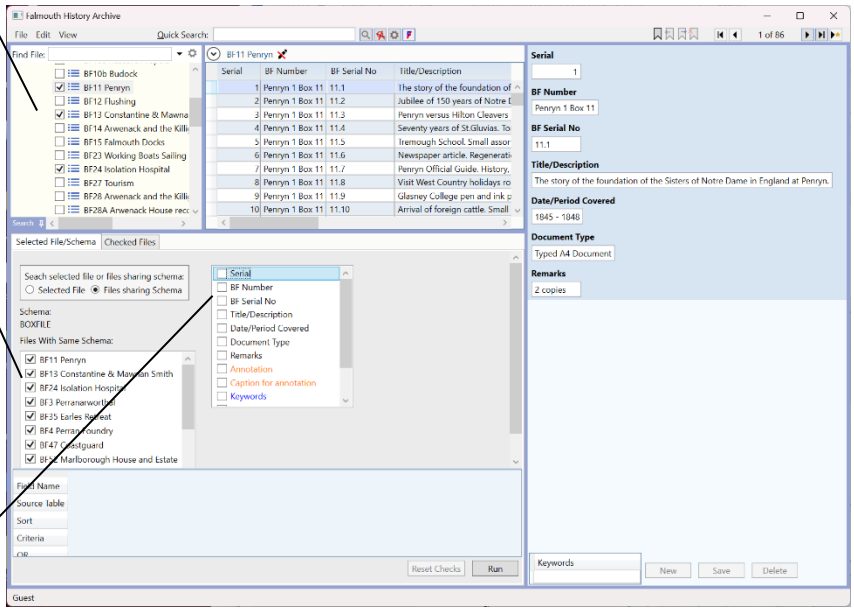
In this example file *BF11 Penryn* has been selected and the search pane displayed. The search defaults to **Selected File**.
Click on **Files sharing Schema** to extend the search to the files listed in the List Box **Files With Same Schema.**



Checked Files have changed to those sharing the same schema.

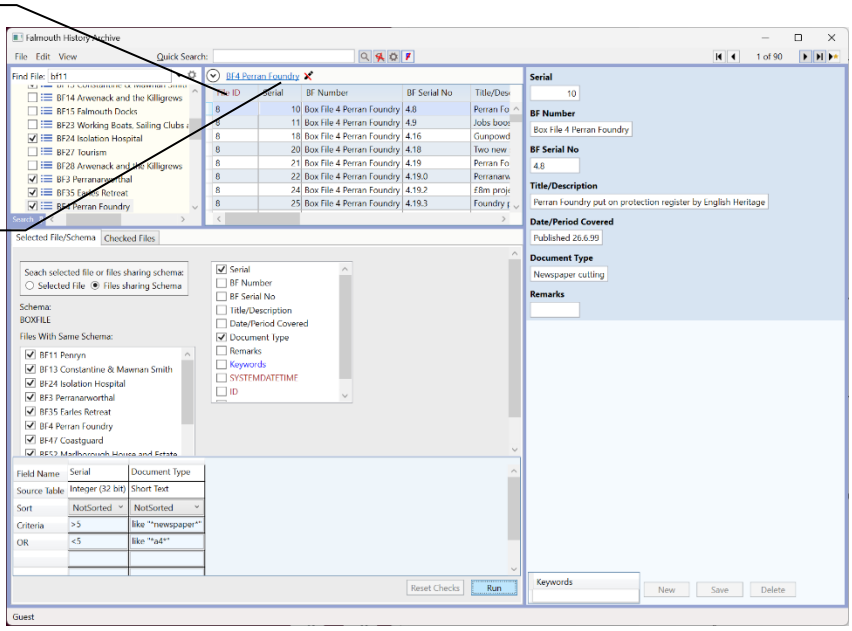
These are the files sharing the same schema. You can check or un-check these files to include or exclude files sharing the same schema

The field list is unchanged from the single file search as these files all have the same schema.



Adding the same criteria as in an earlier example and clicking **Run** returns results from all checked files. The **Data Grid** has an additional **File ID** column so that you can identify returns from differing files…

...and the file name is now a hyperlink that can be clicked to take you to the record within its **Browse** list

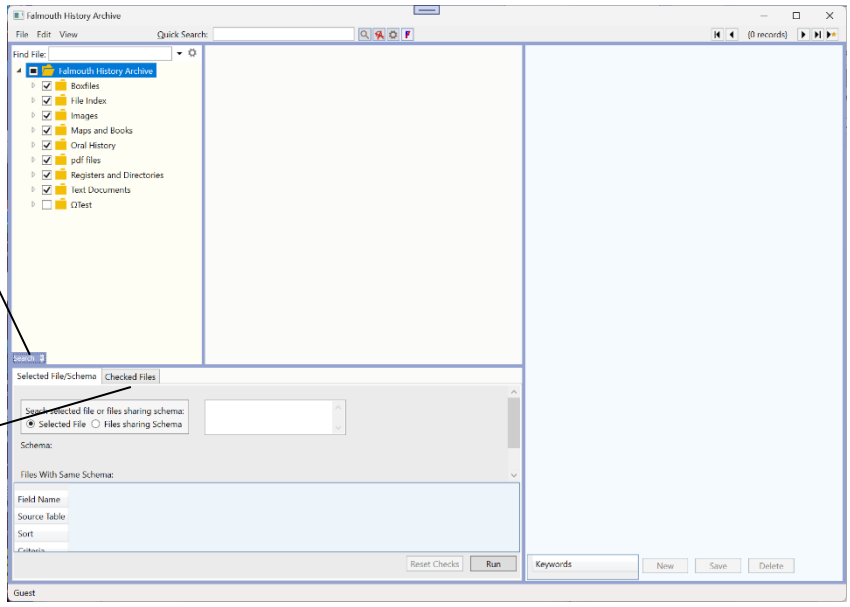


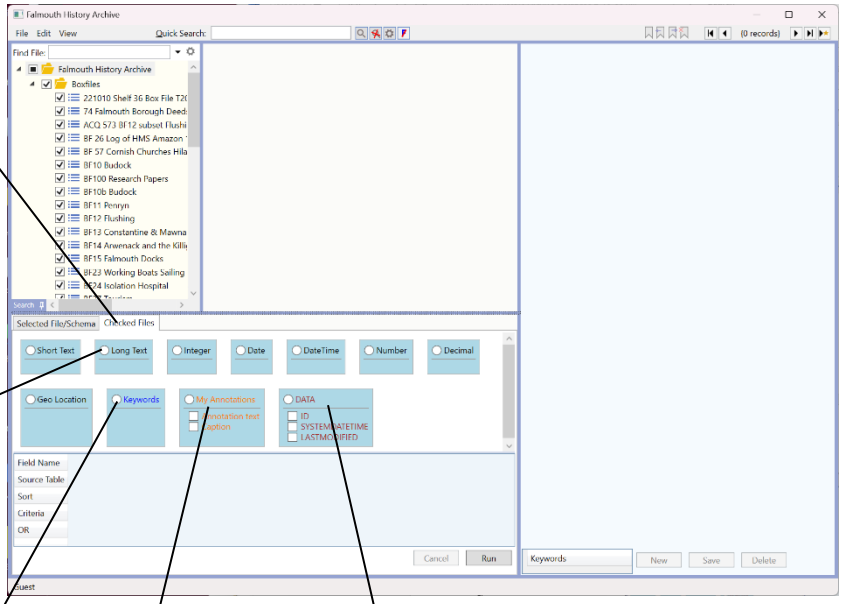
The results of single schema searches can be edited if you are logged in and have edit rights for the respective file.

Searching across multiple schemas with the Checked Files search

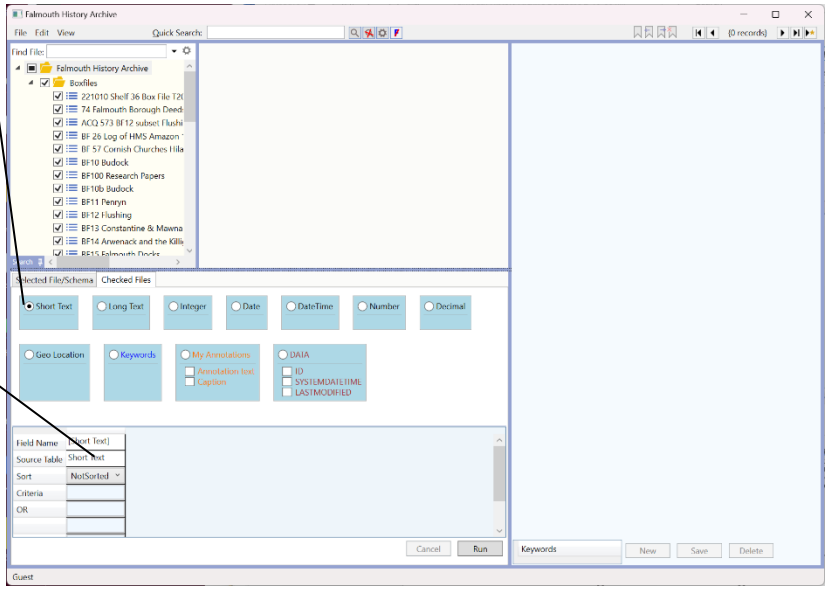
Open the search pane…

…and click the **Checked Files** tab





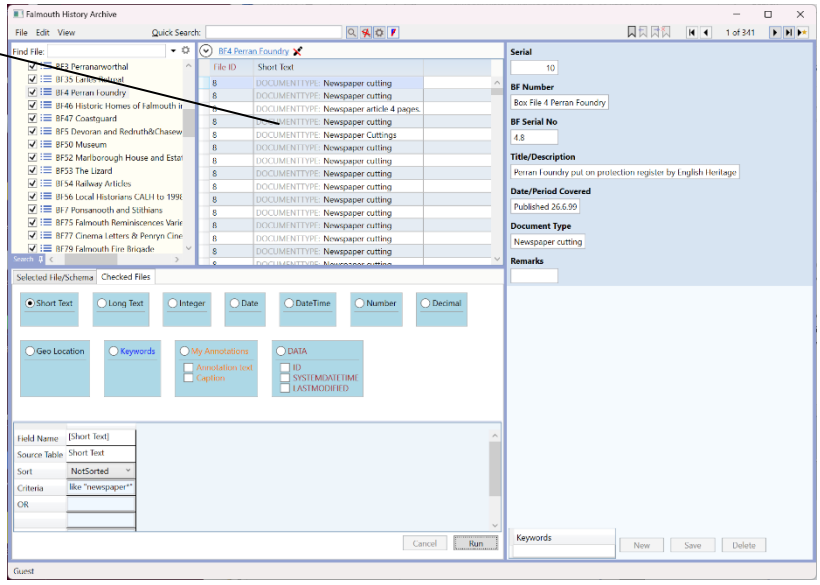
The Checked Files search can only search records by datatype when there are no common fields shared by the checked files.



Add any additional datatypes as required. You can add a datatype without criteria to apply a sort to that column or to view the content of the field. Click **Run** to start the search. The search can be cancelled with the **Cancel** button or by pressing {ESC} on the keyboard.

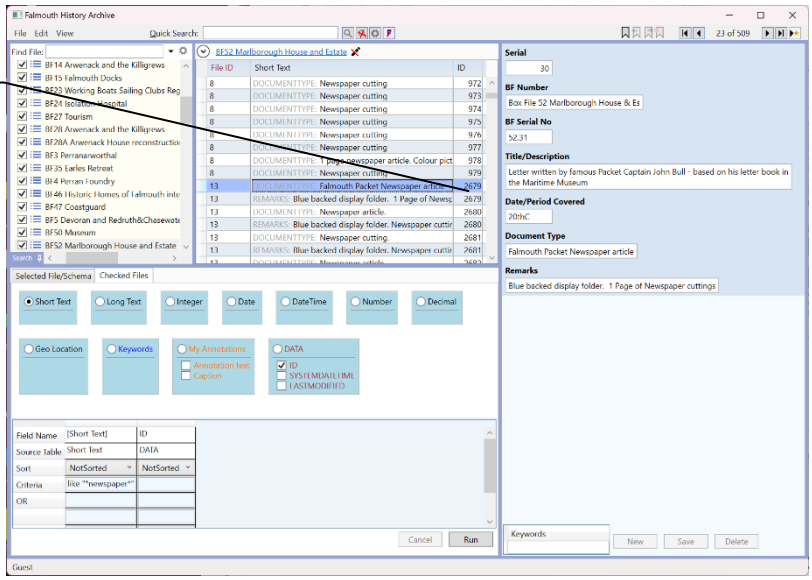
The records returned do not necessarily share the same structure (i.e., Schema) and so the DataGrid output only shows columns for the selected datatypes or fields. Each row in the DataGrid represents a search match or 'find' and, if a record contains more than 1 match, more than 1 row will be returned per record.

Results are shown in the DataGrid. The output only includes the field name in Grey.



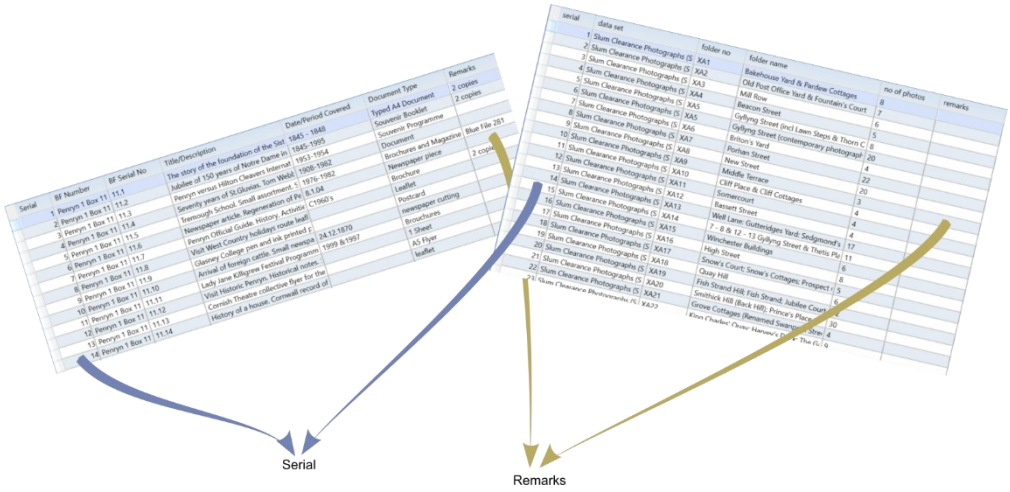
In this example the **Data ID** field has been checked to include that information in the output.

This record has 2 fields of **Datatype Short Text** that contain a match to the criterion. Looking at the ID field output you can see 2 rows from the same record, one from each matched field.

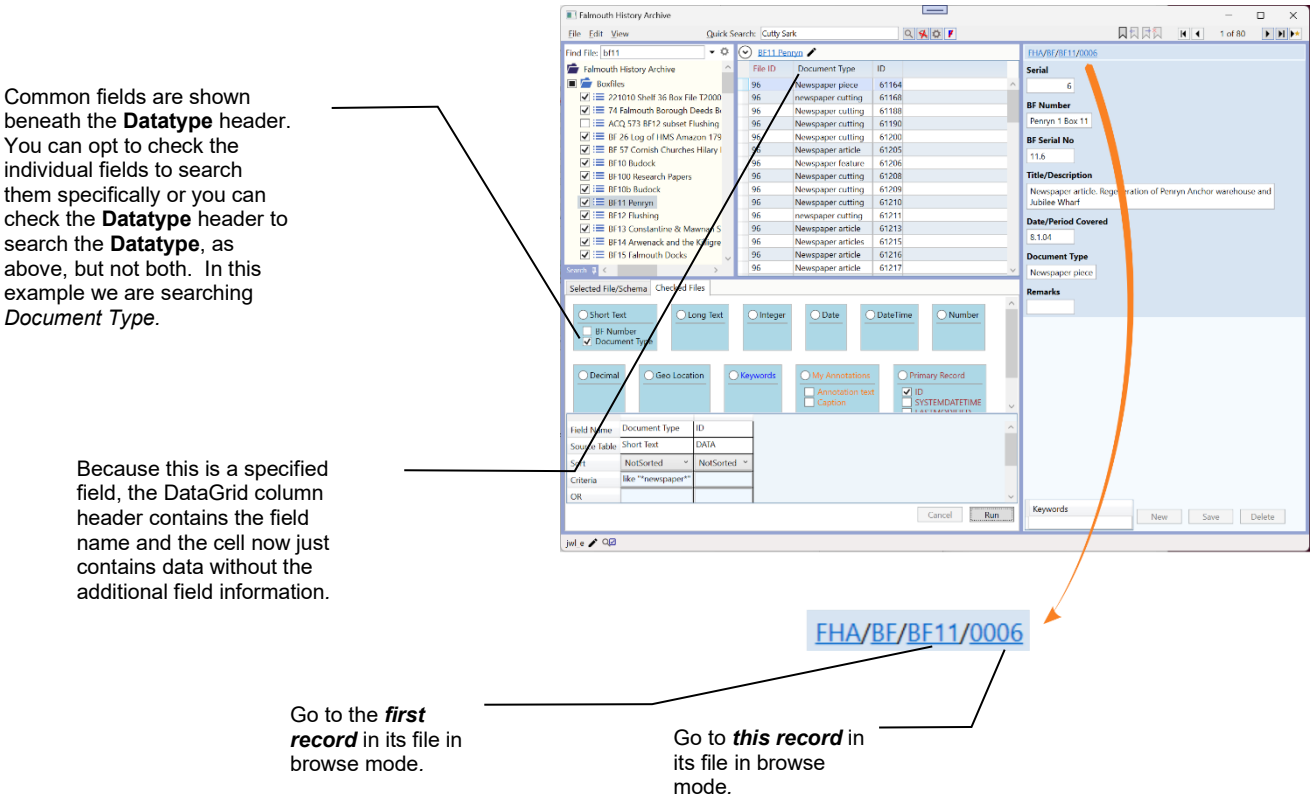


Common fields in different schemas

This illustration shows 2 different files having different schema and consequently different fields. They do, however, have 2 fields in common: ***Serial*** and ***Remarks***. Common fields share the same *name* and *datatype*.



In the example, below, the checked file selection has been reduced to a few Box Files. The selection still spans multiple Schema but the files have 2 common fields: *BF Number* and *Document Type*.



Results of the **Checked Files** search can be edited in the details pane, if you have appropriate file permission, but the DataGrid is read-only.

Overview of ArchiveDb data structure

We have already discussed the Archive Group (or Fonds), Series and File structure of the database. Now we will look at the data structure of individual Files.

Each File stores data in a table consisting of columns and rows, one row per record. Columns are also called **Fields** and the terms are interchangeable. Each Field must have a unique name within the table. As well as a **Field Name**, a column also has a **Caption**. The Caption is the text displayed as the column header. The Field Name is not displayed except in the multi-schema search result grid. Field Names are always uppercase, cannot contain spaces and have a maximum length of 30 characters.

Each Field is assigned one of the following data types:

Data Types

Each Field (or column) has a specific data type. The data types are:

- Short Text
- Long Text
- Date
- Date + Time
- Integer (32 bit)
- Integer (64 bit)
- Decimal (£ etc)
- Number (float)
- Image
- Binary
- Geolocation

Each data type has its own on-disk table, see below.

Short Text

This is the data type most used within ArchiveDb. It holds up to 250 characters of plain Unicode (UTF8[4]) text. These fields are indexed and directly searchable. Words and dates within these fields are saved to the Quick Search indices.

Long Text

This is the second most used data type. It holds plain Unicode (UTF8) text. Each Long Text field can hold a maximum of around 2GB text and can be searched directly. Words and dates within memo fields are saved to the Quick Search indices.

Date Only

Date fields store dates in the format *dd/MM/YYYY*. This means mandatory precision to 1 day. Very often historical data is not that precise. You may only know a year or maybe month and year, in which case the Date type is unsuitable. In this eventuality use a Short Text field. As mentioned above, the parsing process that occurs when a record is saved looks for text that may be a date and stores that information in a separate ArchiveDate index.

[4] UTF8 is a character coding specification

Data held in Date fields is also saved to the ArchiveDate index. The ArchiveDate index stores the whole date, the month and the year as separate entities to allow searching for part dates (i.e., just the year or month and year)

Date + Time

As for Date but includes time (hh:nn[5]:ss) as well. This level of precision is unlikely to be needed and no records in ArchiveDb have so far used this data type.

Integer (32 bit)

Whole numbers from -2^{32} to $2^{32} - 1$, that is, from -2,147,483,648 to 2,147,483,647.

This data type is indexed and directly searchable, and useful for storing serial numbers etc. within individual Files. Integer fields cannot, however, be searched with the Quick Search.

Integer (64 bit)

Whole numbers from -2^{63} to $2^{63} - 1$, or from -9,223,372,036,854,775,808 to 9,223,372,036,854,775,807.

This data type is used internally for record IDs as it can store enormous numbers. It is unlikely that you will need to store such large numbers within the Archive tables and it is not available for user-defined fields.

It is indexed and directly searchable. Integer fields cannot be searched with the Quick Search.

Decimal (£ etc)

This is quite a specialised data type primarily intended for storing currency values. It stores a decimal number with an integer portion 13 digits long and 4 digits after the decimal point. The prefix (£, $ etc) is stored in a separate field within the data type. It provides accuracy in division and multiplication of currency values that will not be much needed within ArchiveDb. The prefix component can be up to 6 characters in length. It might also have a use storing certain coded values, bearing in mind its limitations vs a Short Text field. If you enter a prefix value of EUR, USD or YEN the prefix will change to €, $ and ¥. It is indexed and directly searchable by number and/or prefix. Decimal fields cannot be searched with the Quick Search.

Number (float)

Floating point single precision number e.g. 2.123456 or 12345.6. A number with a decimal point. The more digits before the decimal point, the fewer you can have after it. Assume a maximum precision of 6 or 7 digits in total, adding together those before and after the decimal point.

Image

Stores an image. The field is composite and can store an unlimited number of images per record. Each image page can be stored either as binary data within the database itself or as a separate file with a reference held in the Image data type (preferred). See later.

The image data type cannot be searched.

[5] nn represents minutes in 2 digits

Binary

The binary data type stores binary data (i.e. computer file data). Binary data and image data are stored identically, the only difference is that you will be offered a file select dialog filtered for images if the field is defined as Image but the default is PDF for the binary datatype. It stores computer files such as PDFs, audio files (including oral histories), text documents, Word documents, etc. ArchiveDb has native support for PDF documents and audio-visual files. The text behind PDF documents is indexed and searchable as are any transcripts attached to audio-visual files.

The binary data type cannot be searched.

Geolocation

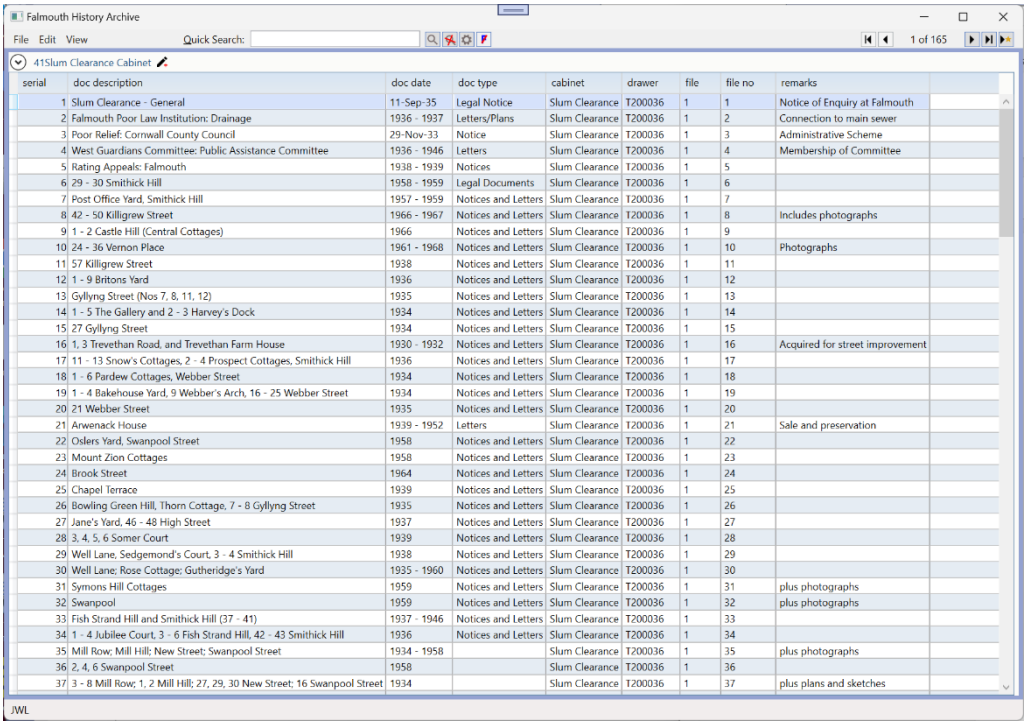
This is a specialised data type that holds latitude and longitude. Internally latitude and longitude are stored in 2 separate double precision floating point fields.

This data is indexed and directly searchable. The search has not been fully implemented but will, in future, provide a search to look for records within a given distance (radius) of the specified point. At the time of writing, ArchiveDb contains no geolocation data. Geolocation fields cannot be searched using the Quick Search.

Table structure

As stated above, each file has a table structure comprised of a series of columns or fields, each with a specified data type, field name and caption. This specification is called the **Schema**.

The image below shows file ***41Slum Clearance Cabinet*** displayed in ArchiveDb. The view menu has been set to show only the Grid View.



serial	doc description	doc date	doc type	cabinet	drawer	file	file no	remarks
1	Slum Clearance - General	11-Sep-35	Legal Notice	Slum Clearance	T200036	1	1	Notice of Enquiry at Falmouth
2	Falmouth Poor Law Institution: Drainage	1936 - 1937	Letters/Plans	Slum Clearance	T200036	1	2	Connection to main sewer
3	Poor Relief: Cornwall County Council	29-Nov-33	Notice	Slum Clearance	T200036	1	3	Administrative Scheme
4	West Guardians Committee: Public Assistance Committee	1936 - 1946	Letters	Slum Clearance	T200036	1	4	Membership of Committee
5	Rating Appeals: Falmouth	1938 - 1939	Notices	Slum Clearance	T200036	1	5	
6	29 - 30 Smithick Hill	1958 - 1959	Legal Documents	Slum Clearance	T200036	1	6	
7	Post Office Yard, Smithick Hill	1957 - 1959	Notices and Letters	Slum Clearance	T200036	1	7	
8	42 - 50 Killigrew Street	1966 - 1967	Notices and Letters	Slum Clearance	T200036	1	8	Includes photographs
9	1 - 2 Castle Hill (Central Cottages)	1966	Notices and Letters	Slum Clearance	T200036	1	9	
10	24 - 36 Vernon Place	1961 - 1968	Notices and Letters	Slum Clearance	T200036	1	10	Photographs
11	57 Killigrew Street	1938	Notices and Letters	Slum Clearance	T200036	1	11	
12	1 - 9 Britons Yard	1936	Notices and Letters	Slum Clearance	T200036	1	12	
13	Gyllyng Street (Nos 7, 8, 11, 12)	1935	Notices and Letters	Slum Clearance	T200036	1	13	
14	1 - 5 The Gallery and 2 - 3 Harvey's Dock	1934	Notices and Letters	Slum Clearance	T200036	1	14	
15	27 Gyllyng Street	1934	Notices and Letters	Slum Clearance	T200036	1	15	
16	1, 3 Trevethan Road, and Trevethan Farm House	1930 - 1932	Notices and Letters	Slum Clearance	T200036	1	16	Acquired for street improvement
17	11 - 13 Snow's Cottages, 2 - 4 Prospect Cottages, Smithick Hill	1936	Notices and Letters	Slum Clearance	T200036	1	17	
18	1 - 6 Pardew Cottages, Webber Street	1934	Notices and Letters	Slum Clearance	T200036	1	18	
19	1 - 4 Bakehouse Yard, 9 Webber's Arch, 16 - 25 Webber Street	1934	Notices and Letters	Slum Clearance	T200036	1	19	
20	21 Webber Street	1935	Notices and Letters	Slum Clearance	T200036	1	20	
21	Arwenack House	1939 - 1952	Letters	Slum Clearance	T200036	1	21	Sale and preservation
22	Oslers Yard, Swanpool Street	1958	Notices and Letters	Slum Clearance	T200036	1	22	
23	Mount Zion Cottages	1958	Notices and Letters	Slum Clearance	T200036	1	23	
24	Brook Street	1964	Notices and Letters	Slum Clearance	T200036	1	24	
25	Chapel Terrace	1939	Notices and Letters	Slum Clearance	T200036	1	25	
26	Bowling Green Hill, Thorn Cottage, 7 - 8 Gyllyng Street	1935	Notices and Letters	Slum Clearance	T200036	1	26	
27	Jane's Yard, 46 - 48 High Street	1937	Notices and Letters	Slum Clearance	T200036	1	27	
28	3, 4, 5, 6 Somer Court	1939	Notices and Letters	Slum Clearance	T200036	1	28	
29	Well Lane, Sedgemond's Court, 3 - 4 Smithick Hill	1938	Notices and Letters	Slum Clearance	T200036	1	29	
30	Well Lane; Rose Cottage; Gutheridge's Yard	1935 - 1960	Notices and Letters	Slum Clearance	T200036	1	30	
31	Symons Hill Cottages	1959	Notices and Letters	Slum Clearance	T200036	1	31	plus photographs
32	Swanpool	1959	Notices and Letters	Slum Clearance	T200036	1	32	plus photographs
33	Fish Strand Hill and Smithick Hill (37 - 41)	1937 - 1946	Notices and Letters	Slum Clearance	T200036	1	33	
34	1 - 4 Jubilee Court, 3 - 6 Fish Strand Hill, 42 - 43 Smithick Hill	1936	Notices and Letters	Slum Clearance	T200036	1	34	
35	Mill Row; Mill Hill; New Street; Swanpool Street	1934 - 1958		Slum Clearance	T200036	1	35	plus photographs
36	2, 4, 6 Swanpool Street	1958		Slum Clearance	T200036	1	36	
37	3 - 8 Mill Row; 1, 2 Mill Hill; 27, 29, 30 New Street; 16 Swanpool Street	1934		Slum Clearance	T200036	1	37	plus plans and sketches

This is the Schema for the table:

Field Name	Caption	Data Type
SYSTEMDATETIME	SYSTEMDATETIME	Date + Time
ID	ID	Integer (64 bit)
FK_FILES	FK_FILES	Integer (64 bit)
ID_PARENT	ID_PARENT	Integer (64 bit)
LASTMODIFIED	LASTMODIFIED	Date + Time
SERIAL	serial	Integer (32 bit)
DOCDESCRIPTION	doc description	Short Text
DOCDATE	doc date	Short Text
DOCTYPE	doc type	Short Text
CABINET	cabinet	Short Text
DRAWER	drawer	Short Text
FILE	file	Short Text
FILENO	file no	Short Text
REMARKS	remarks	Short Text

The first 5 fields, in red, are system fields that are necessary for the operation of the database and should not be changed by the user. The remaining fields, in black, are the user-defined fields visible in the user interface.

All fields in this table except for **FK_FILES** and **ID_PARENT** are searchable.

Each file must have a Schema but schemas can be shared between Files. For example, many of the box files share the same schema and therefore have the same table structure. Sharing schemas makes files cross-searchable.

ArchiveDb tables and on-disk storage

One of the features of ArchiveDb is its ability to cross-search the numerous different table structures of the many Files. When records belonging to a File are saved, the data is not stored on the server in the table structure defined by the File schema but in tables for each data type, plus another for the system fields. Thus, looking at the schema for this table (above) you can see that (of the user-defined columns in black) only 2 data types are used, the **Integer (32 bit)** and the **Short Text**. Therefore, the data for this on-screen table is actually stored on disk in 3 tables, one for the **Integer (32 bit)** data, a second for the **Short Text** data and the third for the system fields (in red).

The File table structure shown on screen is *virtual* and generated by the user interface each time you browse or search for data. On disk, the data is always stored in the same set of tables, no matter how complex the schema. Understanding this will help understand the detailed searches.