AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

Welcome to the ACIC

The Serious Fraud section was created in April 2025 to fulfil the ACIC's contributions in relation to fraud committed against Australian government payments and programs. The ACIC is part of the Fraud Fusion Taskforce (FFT) and hosts the Fraud Fusion Centre (FFC). The FFC is the intelligence lead for the FFT and is the hub for intelligence and data sharing across the FFT.

Fraud Fusion Taskforce

In October 2022, the Government committed $126.3 million over 4 years to establish the FFT, with an aim to reducing the incidents and opportunity for non-compliance and existing or emerging serious and organised crime (SOC) exploitation of the National Disability Insurance Scheme (NDIS) and other Government payments and programs. The FFT consists of the following 21 member agencies and is jointly led by the National Disability Insurance Agency (NDIA) and Services Australia.

- Aged Care Quality and Safeguards Commission
- Attorney-General's Department – Commonwealth Fraud Prevention Centre
- Australian Charities and Non for Profits Commission
- Australian Securities and Investments Commission
- Australian Skills Quality Authority
- Australian Criminal Intelligence Commission
- Australian Federal Police
- Australian Taxation Office
- Australian Transaction Reports and Analysis Centre (AUSTRAC)
- Tax Practitioners Board
- Commonwealth Director of Public Prosecutions
- Department of Employment and Workplace Relations
- Department of Veterans' Affairs
- Department of Education
- Department of Health and Aged Care
- Department of Social Services
- National Disability Insurance Agency
- National Indigenous Australians Agency
- NDIS Quality and Safeguards Commission
- Professional Services Review Agency
- Services Australia.

Fraud Fusion Centre

The Australian Criminal Intelligence Commission (ACIC) leads the Fraud Fusion Centre (FFC), consisting of representatives from the ACIC and some FFT member agencies. The FFC operates as the primary mechanism to enable data acquisition, detection analytics, and intelligence generation and sharing between participating FFT agencies in the FFC. The Fraud Fusion Centre is made up of a core group of ACIC employees and staff from Taskforce agencies embedded in the ACIC. The Fraud Fusion Centre is based in

Canberra but has staff operating out of the s 47E(d) The overall strategic aim of the FFC is to enhance the national understanding existing and emerging SOC fraud against Government payments and programs, and deliver risk indicator and treatment options based on the shared intelligence of FFC. s 37(2)(b), s 47E(d)

s 47E(d)

More information can be found on ACIC Intranet page: Serious Financial Crime | Your Intranet

Team Structure

s 47E(d)

AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

My Learning

Online learning modules to complete (MyLearning URL: Home (acic.gov.au)). If these haven't been completed, you will need to then enrol and complete the below courses:

s 22(1)(a)(ii)

- Getting Started with Palantir Gotham

s 22(1)(a)(ii)

AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

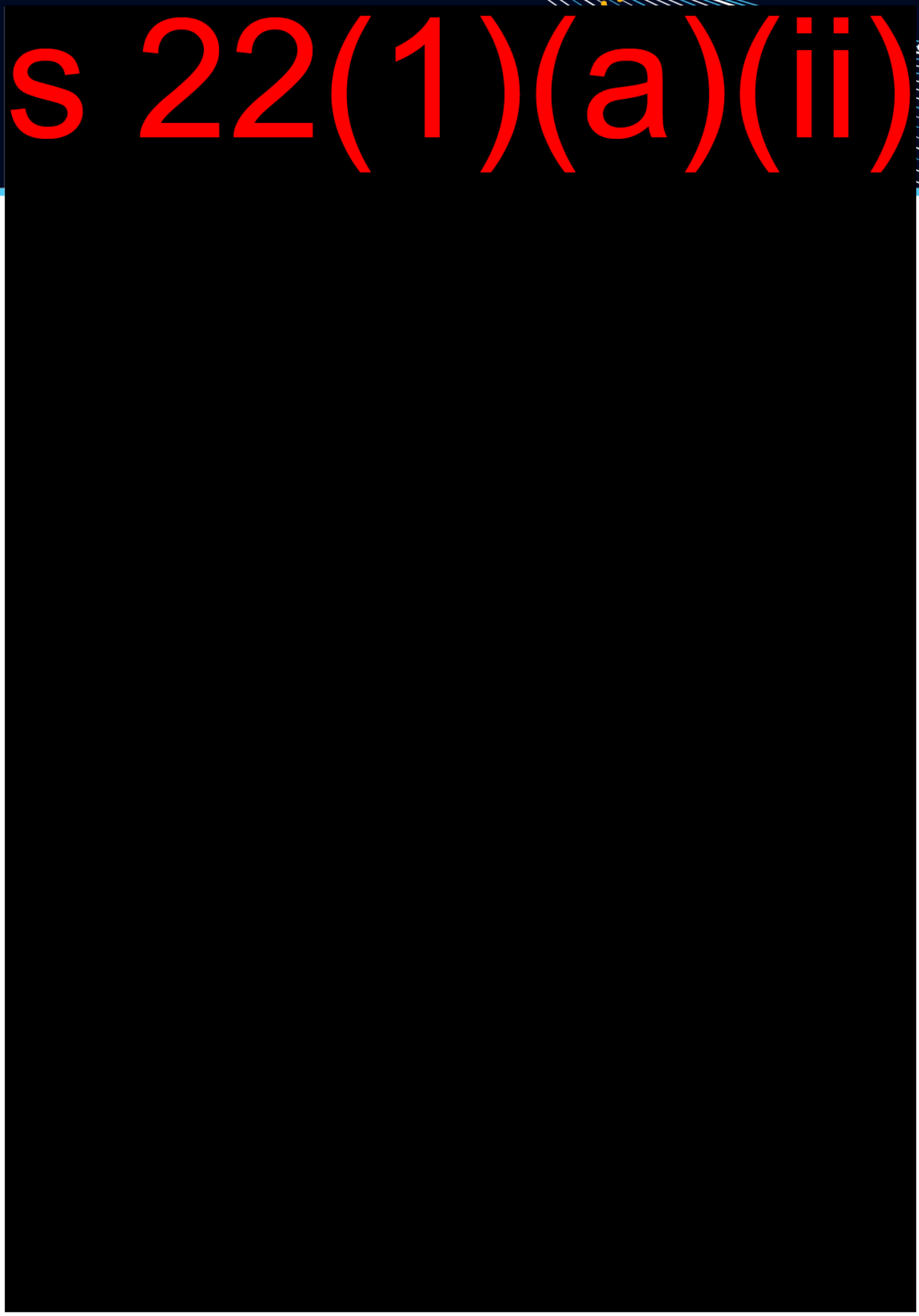
Quick guide to requesting ACIC systems access for new starters in FFC

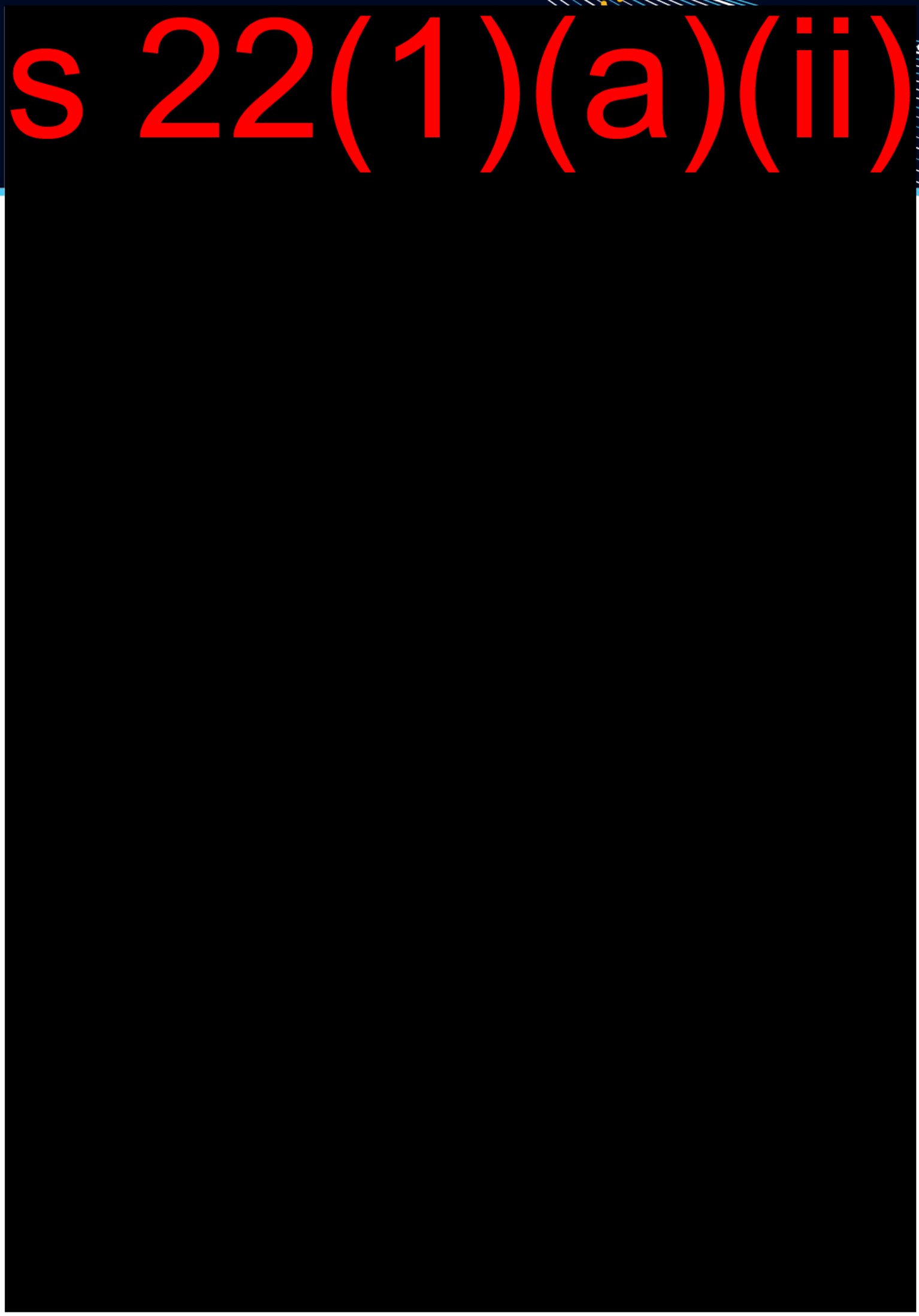
s 22(1)(a)(ii)

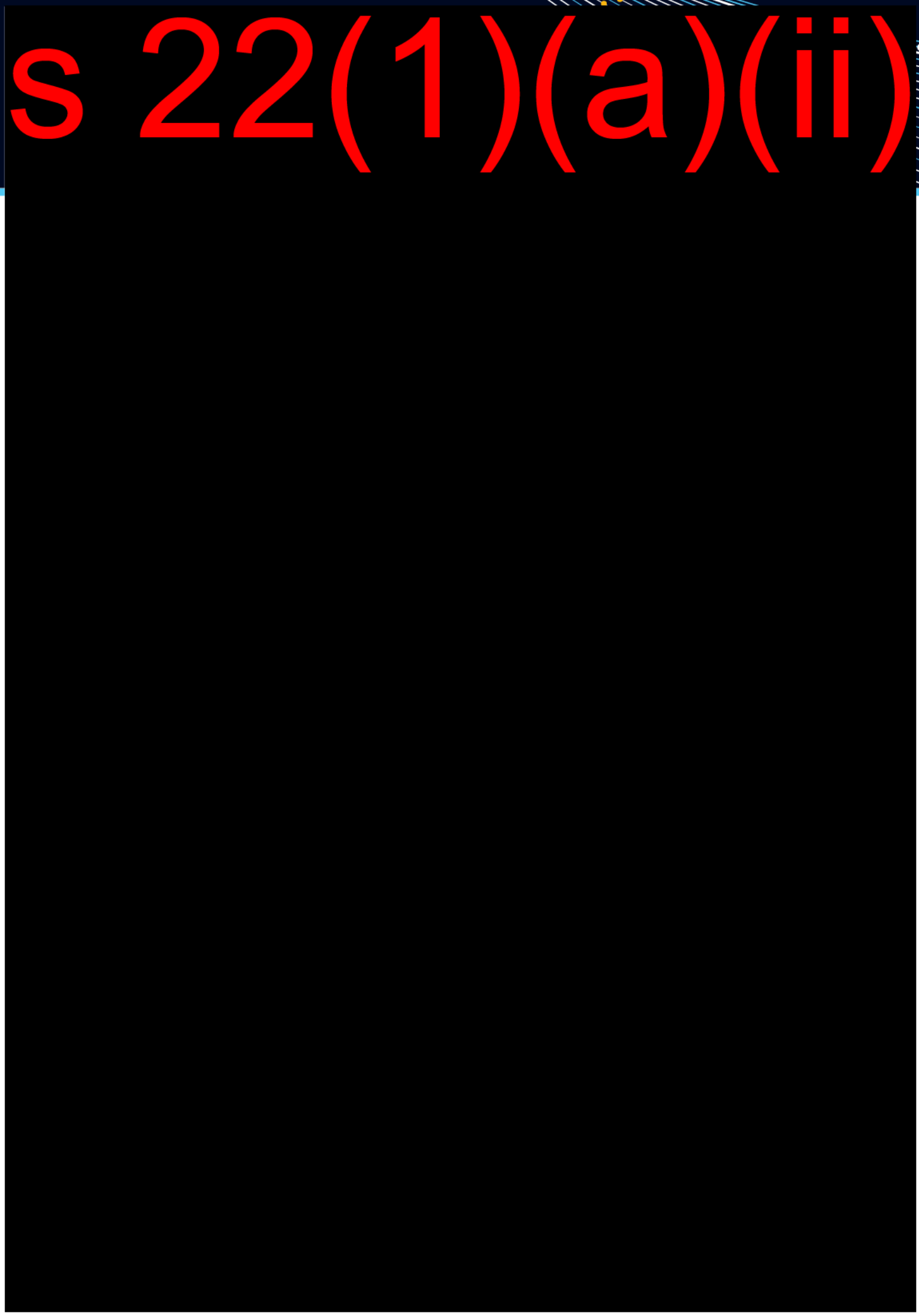
AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

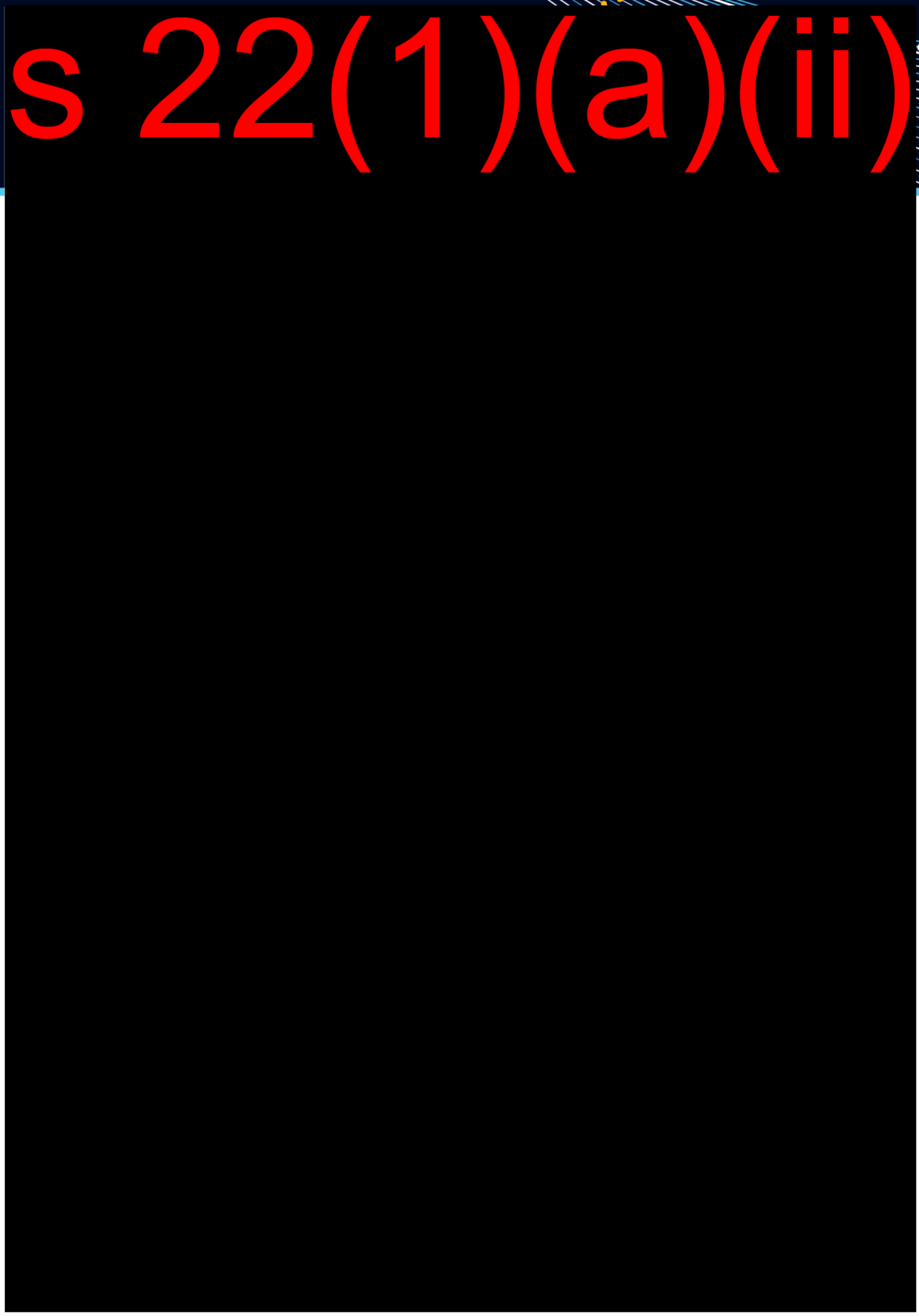
Palantir	ASK request (EL 2 approval required) Note that if the staff member is going to be doing s 47E(d) uploads and analysis, **Power User access** must be requested.	Requested ☐ Access granted ☐

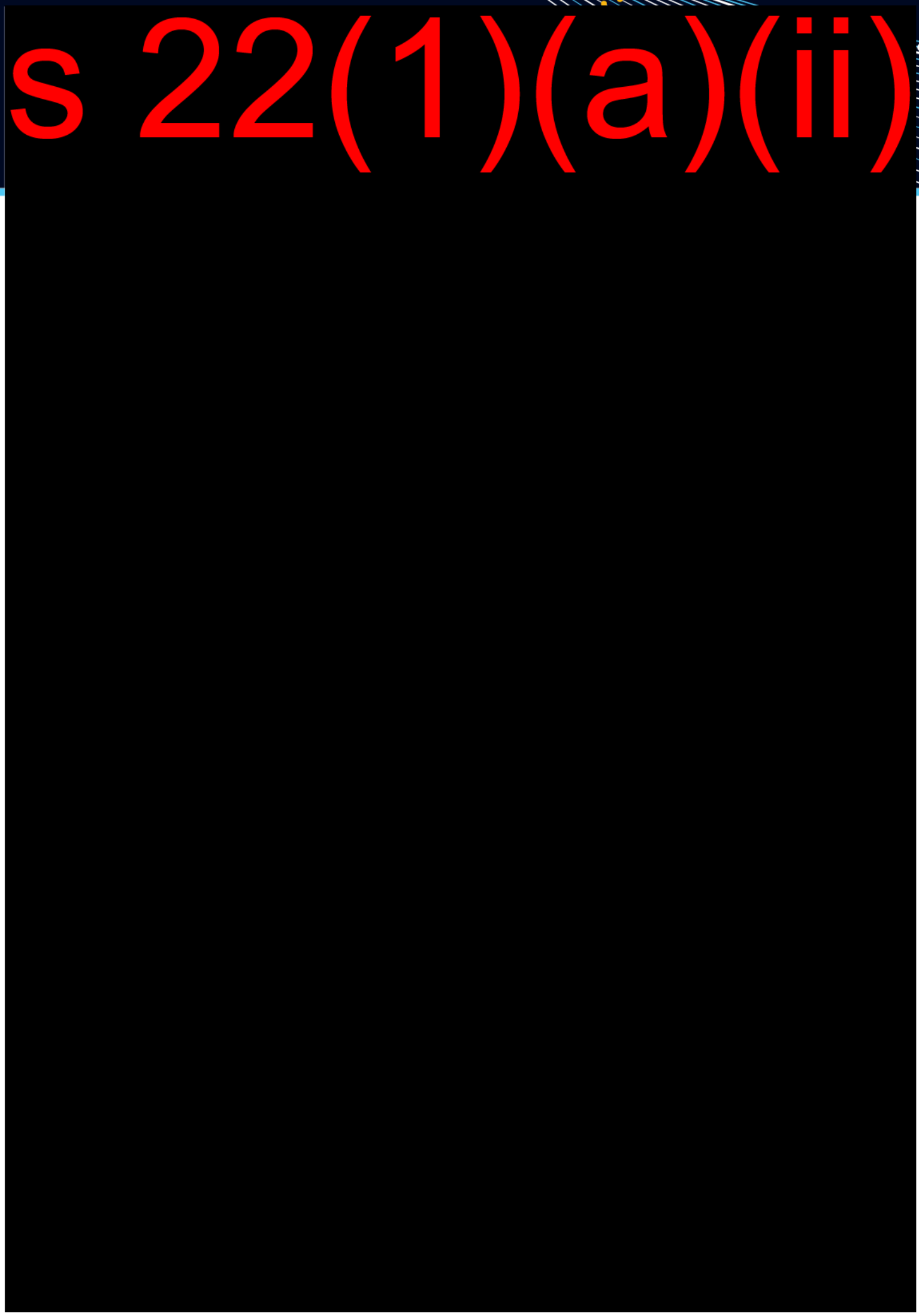
s 22(1)(a)(ii)


s 22(1)(a)(ii)


s 22(1)(a)(ii)


s 22(1)(a)(ii)


s 22(1)(a)(ii)


s 22(1)(a)(ii)

USER GUIDE – FFC INTELLIGENCE COLLECTION (SYSTEMS & DATA SETS)

Updated April 2026

INTRODUCTION:

In October 2022, the Government establish the Fraud Fusion Taskforce (FFT), with an aim of reducing the incidents and opportunity for non-compliance and existing or emerging serious and organised crime (SOC) exploitation of the National Disability Insurance Scheme (NDIS) and other Government payments and programs. The FFT consists of 24 member agencies and is jointly led by the National Disability Insurance Agency (NDIA) and Services Australia.

The Australian Criminal Intelligence Commission (ACIC) leads the Fraud Fusion Centre (FFC), consisting of representatives from the ACIC and some FFT member agencies. The FFC operates as the primary mechanism to enable data acquisition, detection analytics, and intelligence generation and sharing between participating FFT agencies in the FFC.

ACIC have various systems and data sets to collect intelligence against, however some access is restricted to licensing availability or role. The use of these systems must be for the purposes of the ACIC project you are working on and are subject to use restriction rules. A list of FFC staff access to particular systems can be found here: s 47E(d)

s 37(2)(b), s 47E(d)

s 47E(d)

Guide to using Systems for FFC

Refer to Annexure A for list of the various law enforcement and intelligence systems available to ACIC staff, what they are used for and system location.

Each system will have different use restrictions and ability to on-disclose. When using relevant information in intelligence products from these systems, a judgement needs to be made on the level of detail included, depending on relevance, classification, source of information, <….>. Information obtained from these systems should not be a word for word replication of what you are seeing.



s 22(1)(a)(ii)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

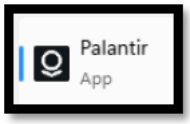
s 22(1)(a)(ii)

PALANTIR

System User Guide: Palantir Gotham Workspace — Gotham Workspace 3.20 documentation

ACIC Training Guides: s 47E(d)

Ensure the "Palantir" app is opened s 37(2)(b), s 47E(d) when opening the program on your desktop. s 37(2)(b), s 47E(d) is for training purposes.



There are two parts to Palantir that are searchable – s 37(2)(b), s 47E(d) and s 47E(d), s 37(2)(b)

s 37(2)(b), s 47E(d)

PALANTIR

s 37(2)(b), s 47E(d)

s 37(2)(b), s 47E(d) **Application**

s 37(2)(b), s 47E(d) is an application that allows you to search s 47E(d) and import these results into your investigation. The s 47E(d) are a collection of large datasets that sit outside s 47E(d) As large datasets can be complex to search, s 47E(d) has various templates to narrow down your results, such as person, address and phone number search templates. All results from s 47E(d) application will be ranked in order of how well they match your search criteria.

Refer to the below guide for instruction on how to use s 47E(d) search function:

s 47E(d)

FFC Note: Source of information needs to be referenced in the endnotes of a product, referencing "Palantir holdings" as this source is not acceptable.

Use Restrictions: Source data set or reporting is required to be identified with the source information on-disclosure rules being applicable.

Legal consideration: Refer to FFC Data Project Confluence Page for more information regarding FFT dataset use - s 47E(d)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

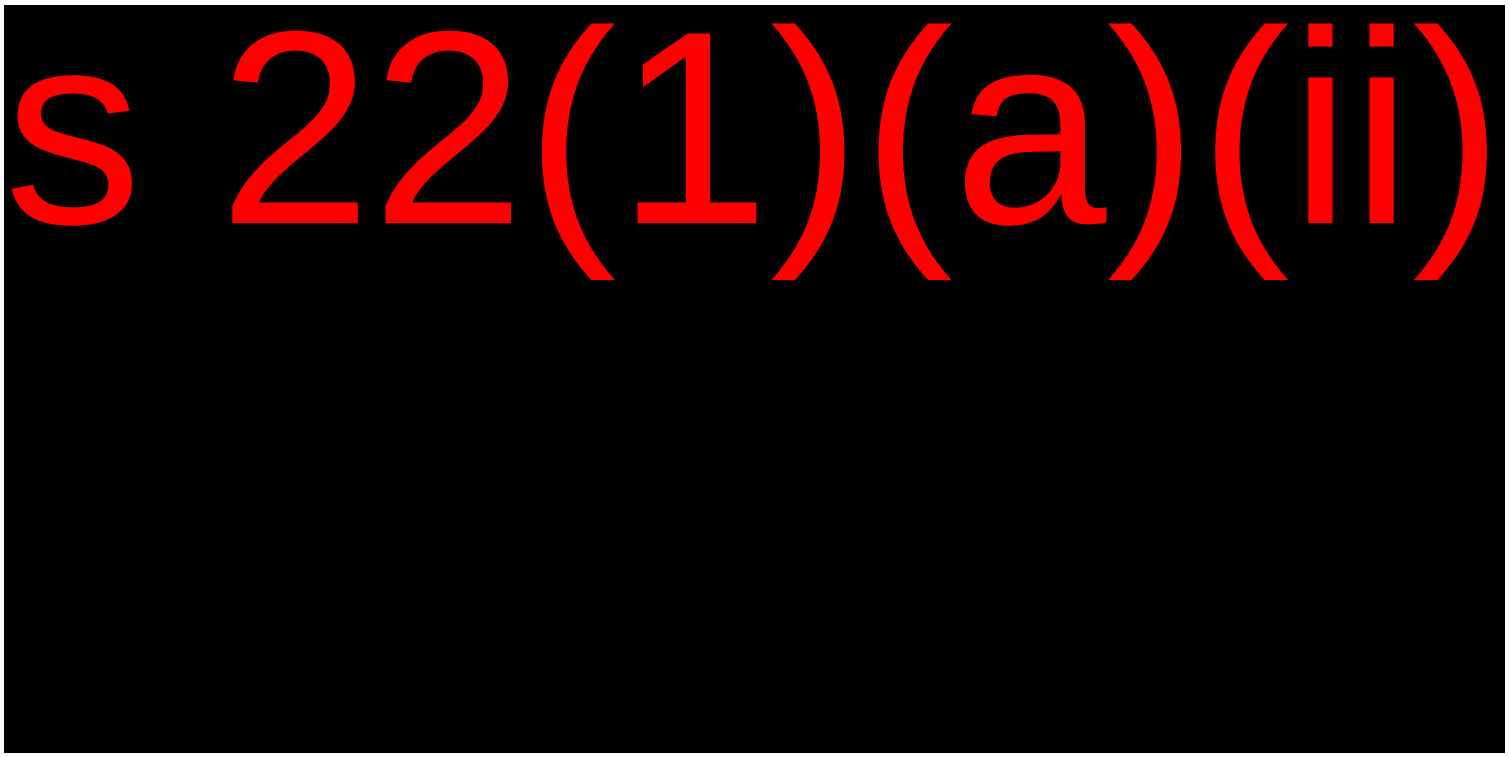
s 22(1)(a)(ii)

FFT Data Sets

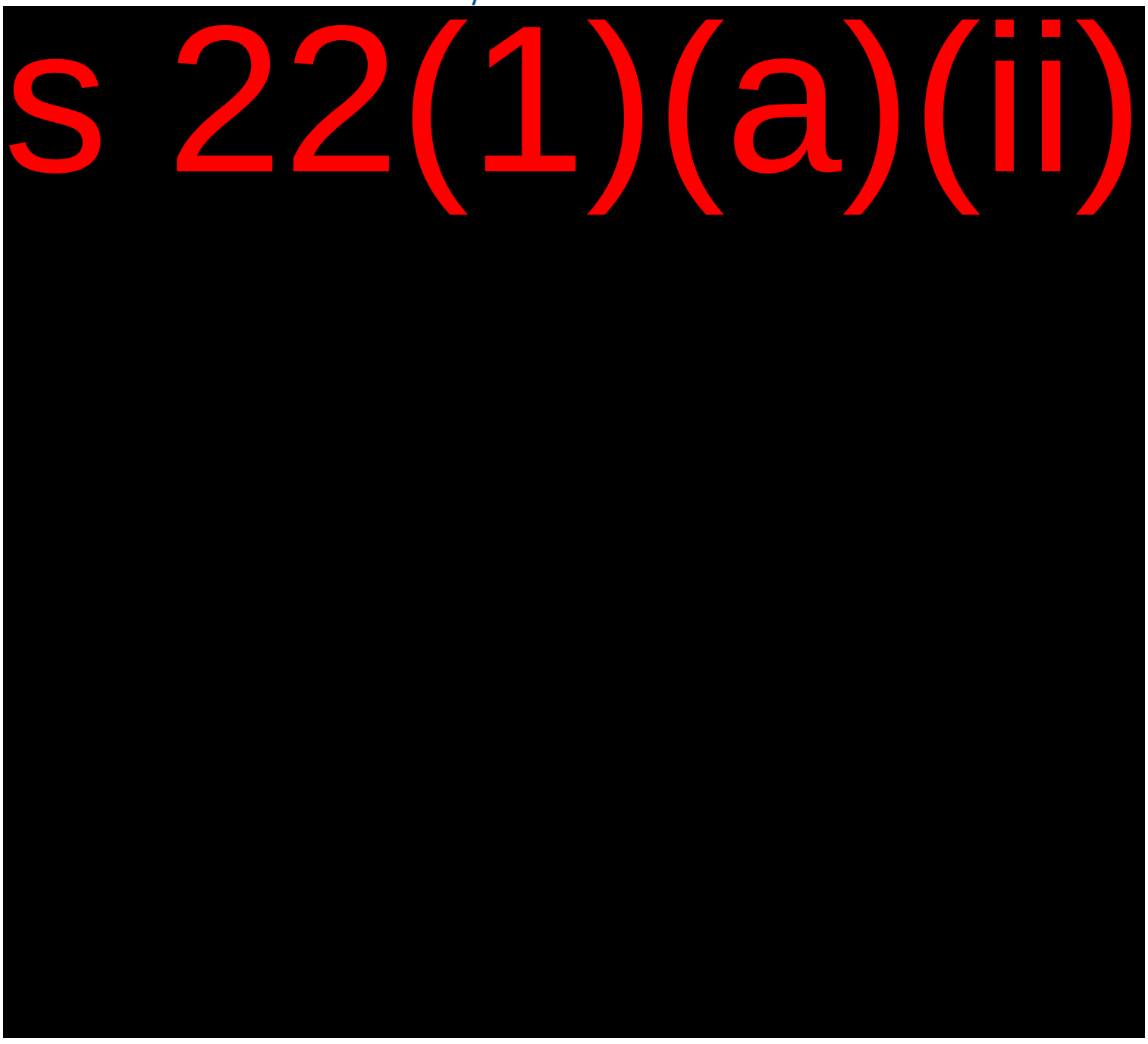
s 47E(d), s 37(2)(b)

2(1)(a)(ii)

s 22(1)(a)(ii)



Annexure A – List of ACIC Systems



SYSTEM	USE	ACCESS
s 22(1)(a)(ii)		
Palantir	Manual entry network mapping and s 47E(d)	Desktop Application
s 22(1)(a)(ii)		

OFFICIAL

Palantir User Guide – Searching Palantir Data Holdings

This Palantir user guide will provide an overview of **the methods and tools you can use when searching Palantir Data Holdings,** covering basic to advanced searching tips for beginners and experienced users. The following tasks will be covered:

1. Palantir s 47E(d) search
   - Homepage search
   - Quick search
   - Keyword search
2. Using the s 47E(d) application
   - s 47E(d)
   - s 22(1)(a)(ii)
   -
3. Advanced searches in Palantir
   - Boolean operators
   - Exact phrases
   - Wildcards
   - Fuzzy operators
4. Search Around tool
5. Customised searches in Palantir
6. Object Explorer

Intelligence Training Contacts

For any feedback or suggestions regarding the contents of this SOP please contact the Intelligence Training team on the contact details below:

- **Email**: s 47E(d)
- **Phone:** s 47E(d)
- **Senior Systems Trainer**: s 22(1)(a)(ii)

Frequently Asked Questions (FAQ's)

What search features are available in Palantir?

- Basic searches can be completed through the homepage, quick search bar and the browser application
- Advanced search operations are available in Palantir, you can use Boolean, wildcard and fuzzy search operators to refine your results (see *Step 3: Advanced Searches in Palantir*)
- The s 47E(d), s 37(2)(b) application can be used to search several different data repositories including the s 47E(d) ,s 22(1)(a)(ii)
- The search around tool can be used on the graph application to search the Palantir s 47E(d), s 37(2)(b) for related objects and property types (*see Step 5: Search Around Tool*)

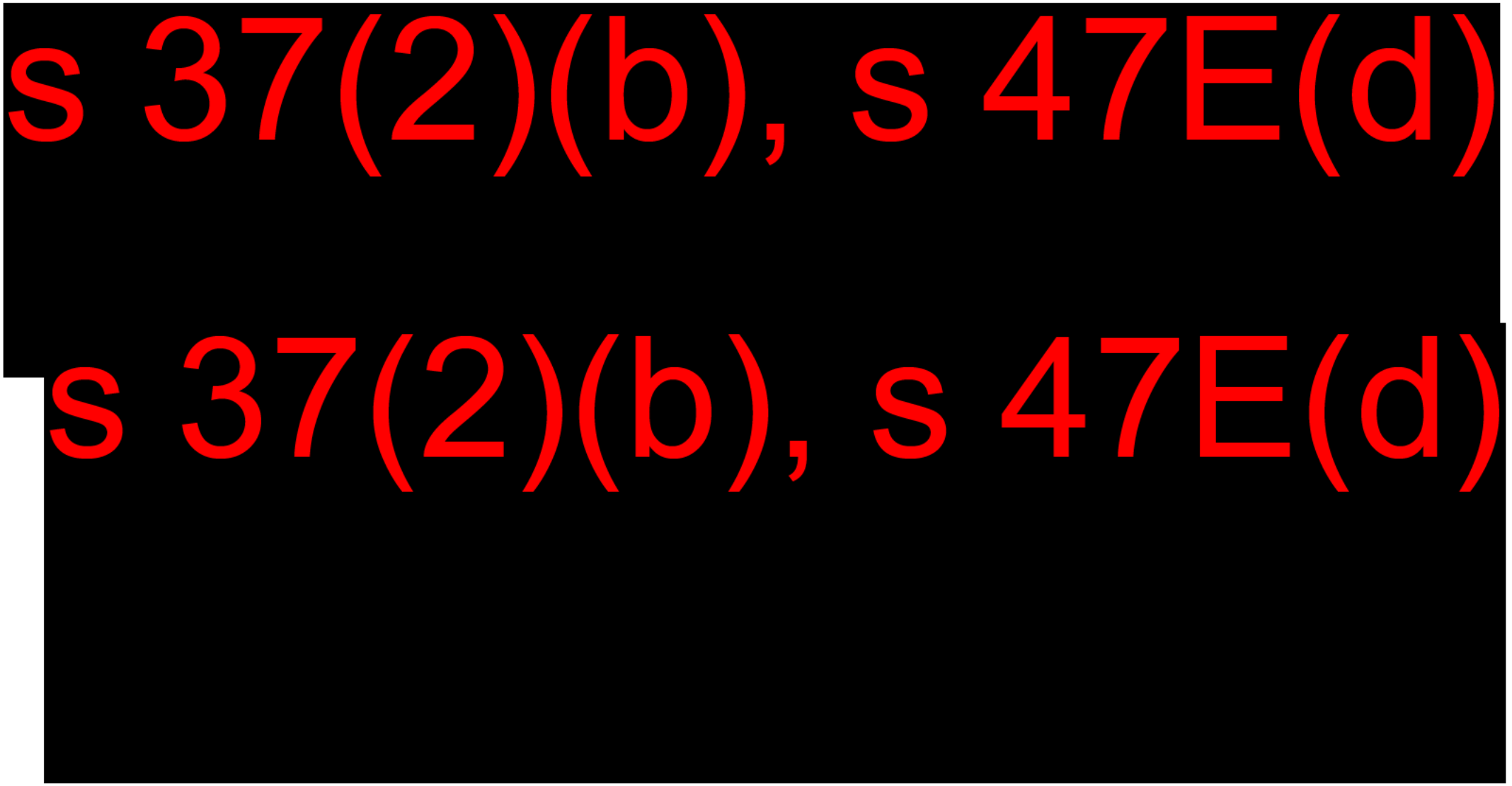
What is the Palantir s 47E(d), s 37(2)(b)

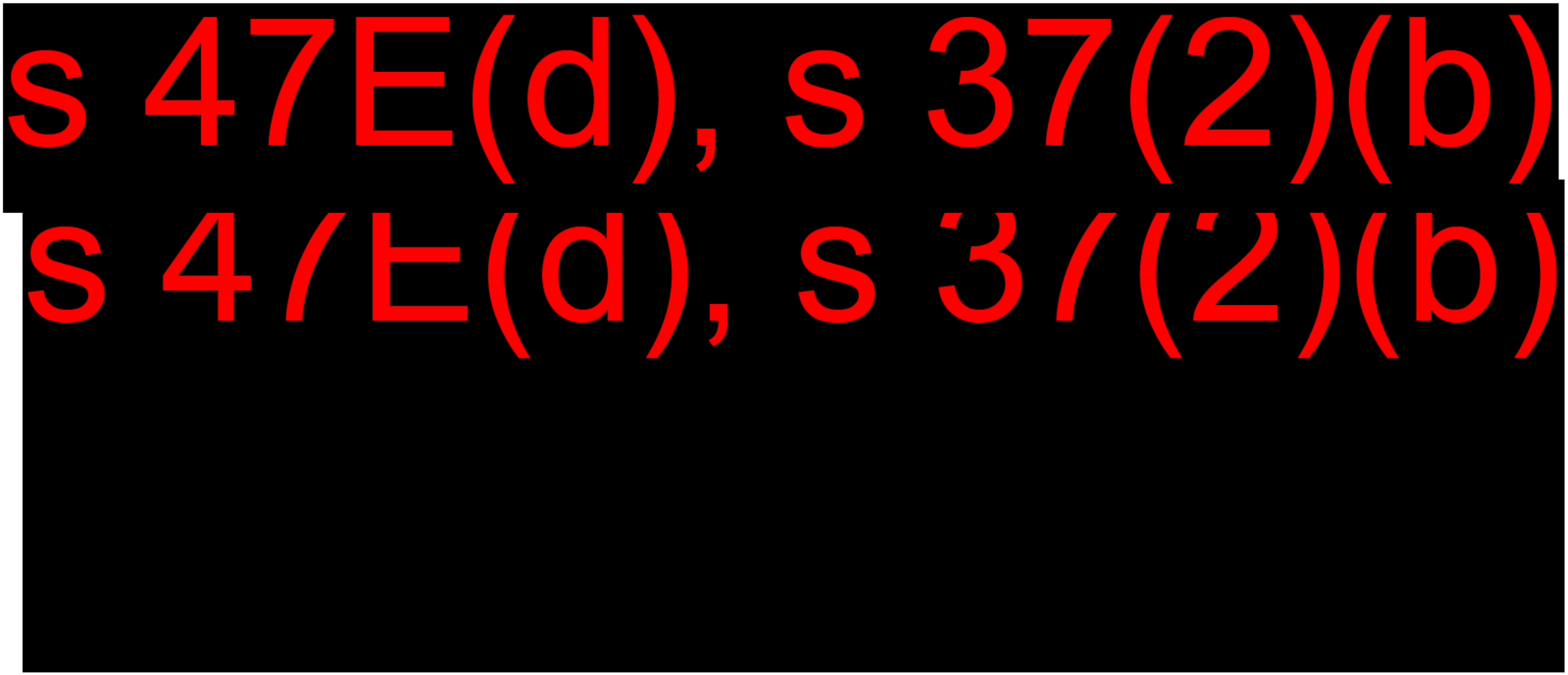
What is the s 47E(d), s 37(2)(b)?

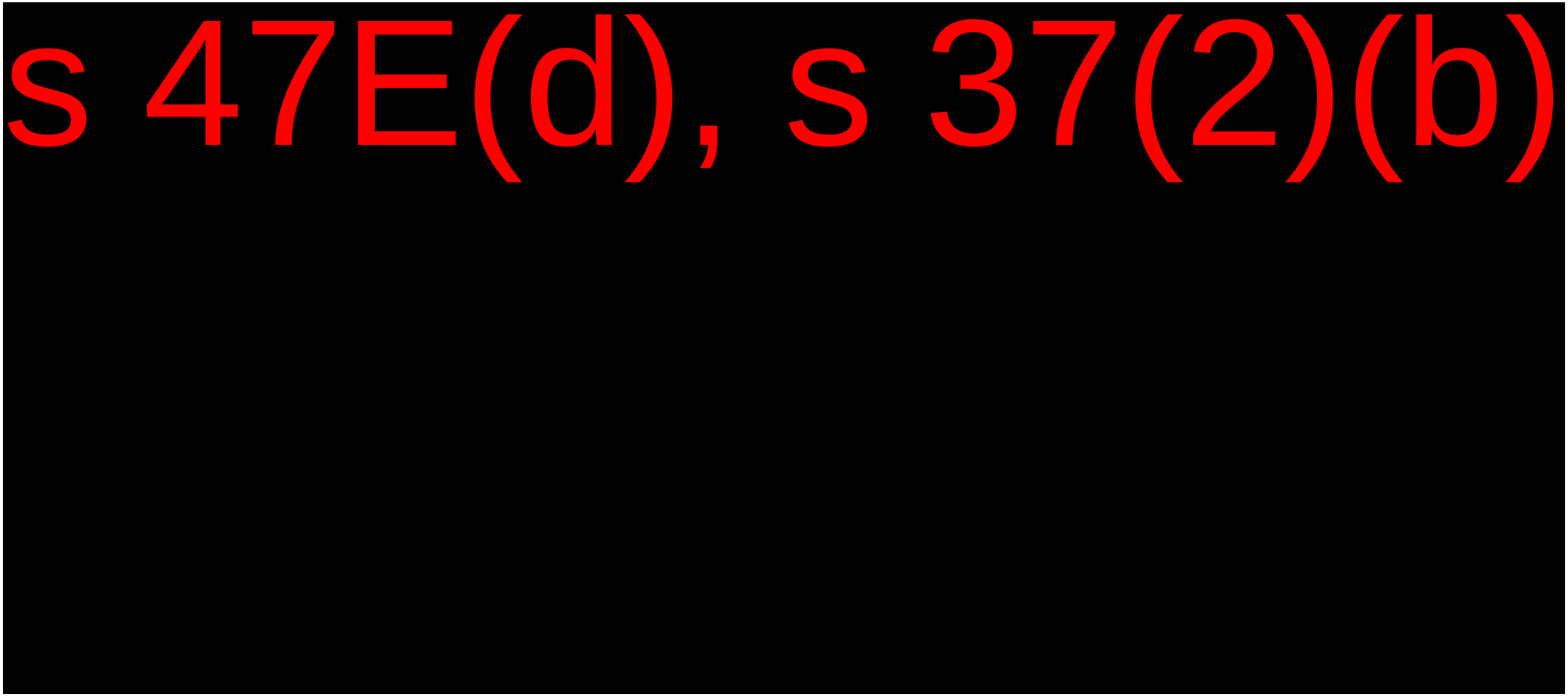
- s 37(2)(b), s 47E(d)
- s 47E(d), s 37(2)(b)

Other Considerations

- s 47E(d), s 37(2)(b)


s 37(2)(b), s 47E(d)
s 37(2)(b), s 47E(d)

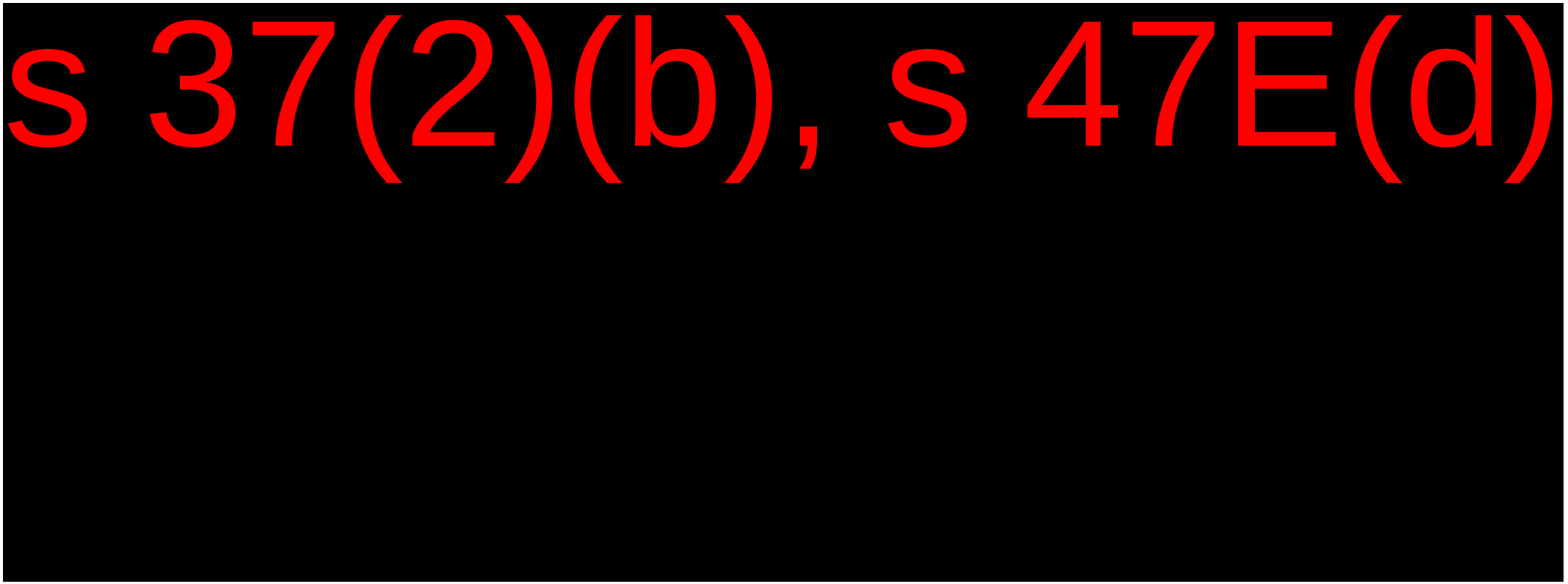

s 47E(d), s 37(2)(b)


s 47E(d), s 37(2)(b)

Quick Search

1. The **quick search bar** is located at the **top-right of the homepage**. This searches the Palantir s 37(2)(b), s 47E(d) and all results will be displayed in a drop-down list. You can select objects directly from this list or open all the results in the browser application

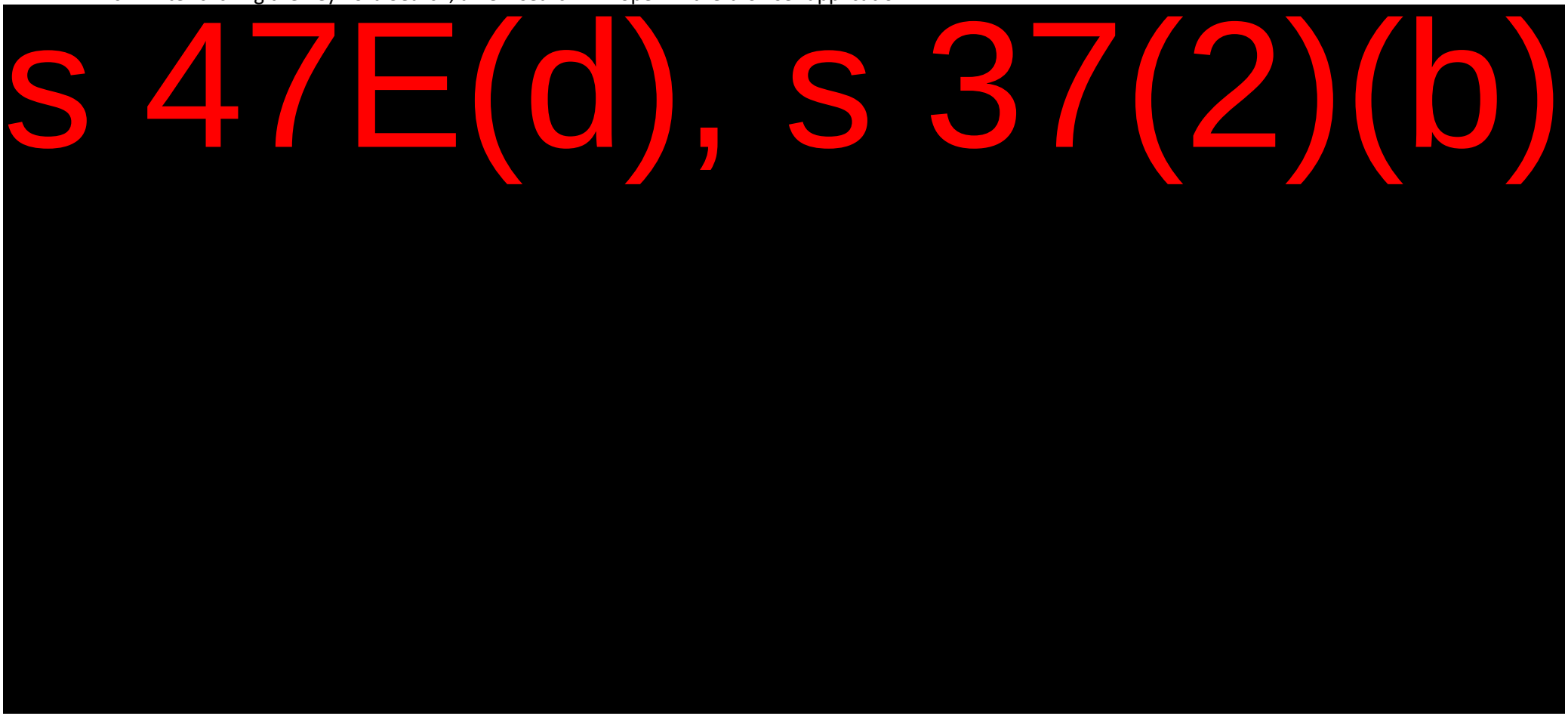
Please note: you can use the **Fuzzy** and **Wildcard** advanced search operators in the quick search bar *(see Step 3: Advanced Searches in Palantir for further information on these operators)*.



Keyword Search

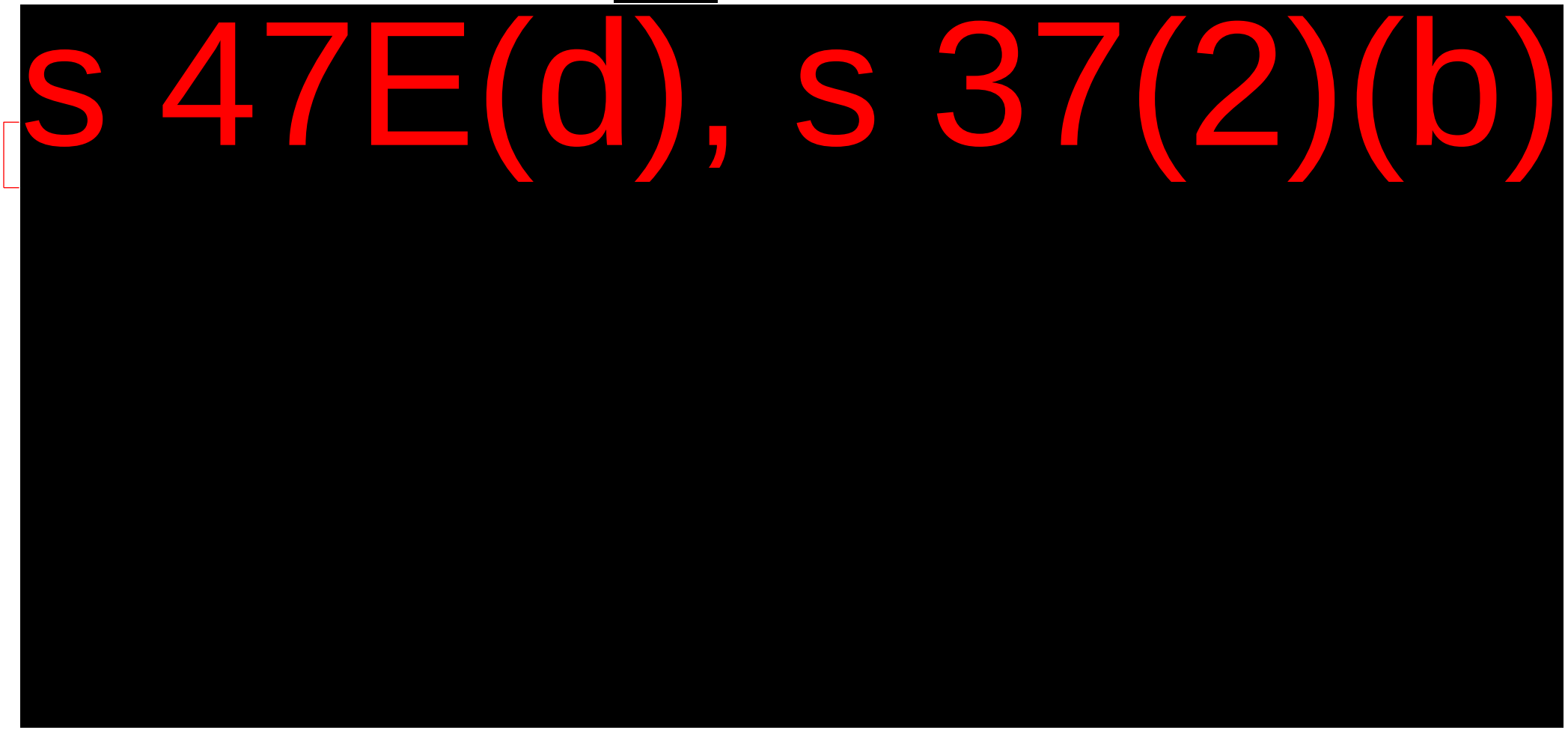
The keyword search is accessed on the homepage below the search bar in Palantir [redacted] In Palantir [redacted] this keyword search function is enabled by clicking on the 'Advanced' hyperlink directly right of the homepage search bar.

1. Click the '**Keyword Search**' button below the homepage search bar
    - After clicking the Keyword Search, a new search will open in the browser application.



2. The keyword search template will allow you to apply filters, such as **dates**, **object types** and **properties.** You can use advanced search operators in this search template as well *(see Step 3: Advanced Searches in Palantir)*
3. In the below instance, we have applied filters for **object type (person)** and **property type (gender = male).** s 47E(d), s 37(2)(b)

Please note: this search template will look different in Palantir s 47E(d), s 37(2)(b) but it ultimately operates the same

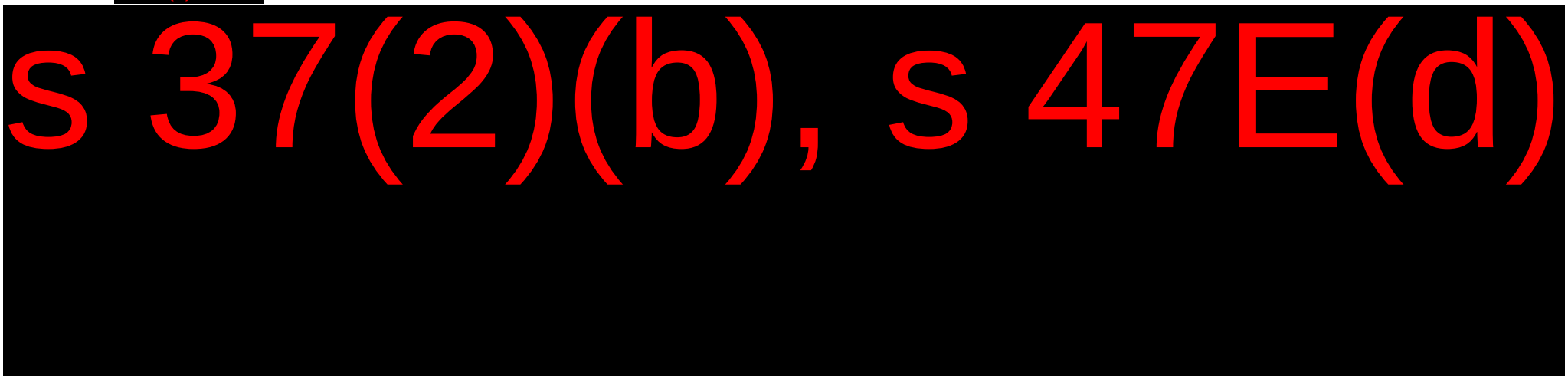


Step 2: Using the s 47E(d), s 37(2)(b) Application

The s 47E(d), s 37(2)(b) is an application that allows you to search s 47E(d) s 22(1)(a)(ii) import these results into your investigation. The s 47E(d) are a collection of large datasets that sit outside the Palantir s 37(2)(b), s 47E(d) As large datasets can be complex to search, s 47E(d), s 37(2)(b) has various templates to narrow down your results, such as person, address and phone number search templates. All results from s 37(2)(b), s 47E(d) application will be ranked in order of how well they match your search criteria.

Searching the s 47E(d)

1. The s 47E(d) application can be enabled from the 'Application' toolbar (first time users will need to do this)

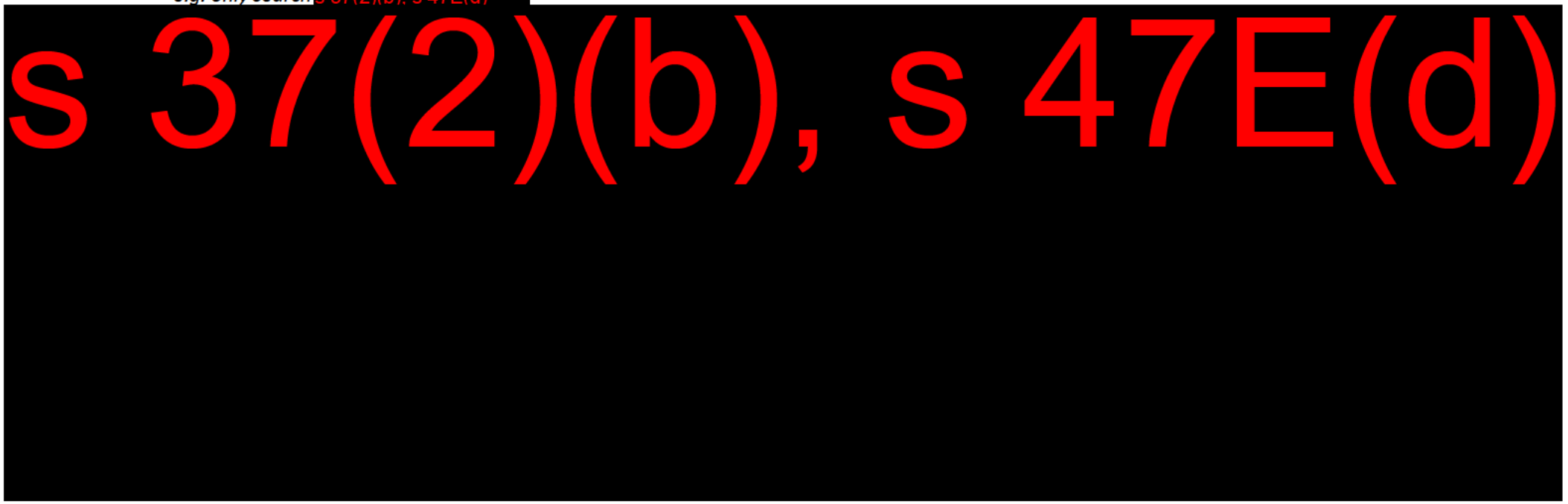


2. If using s 47E(d), s 37(2)(b) for the first time you will need to s 47E(d) to access the [redacted]
    a. s 47E(d) [redacted]

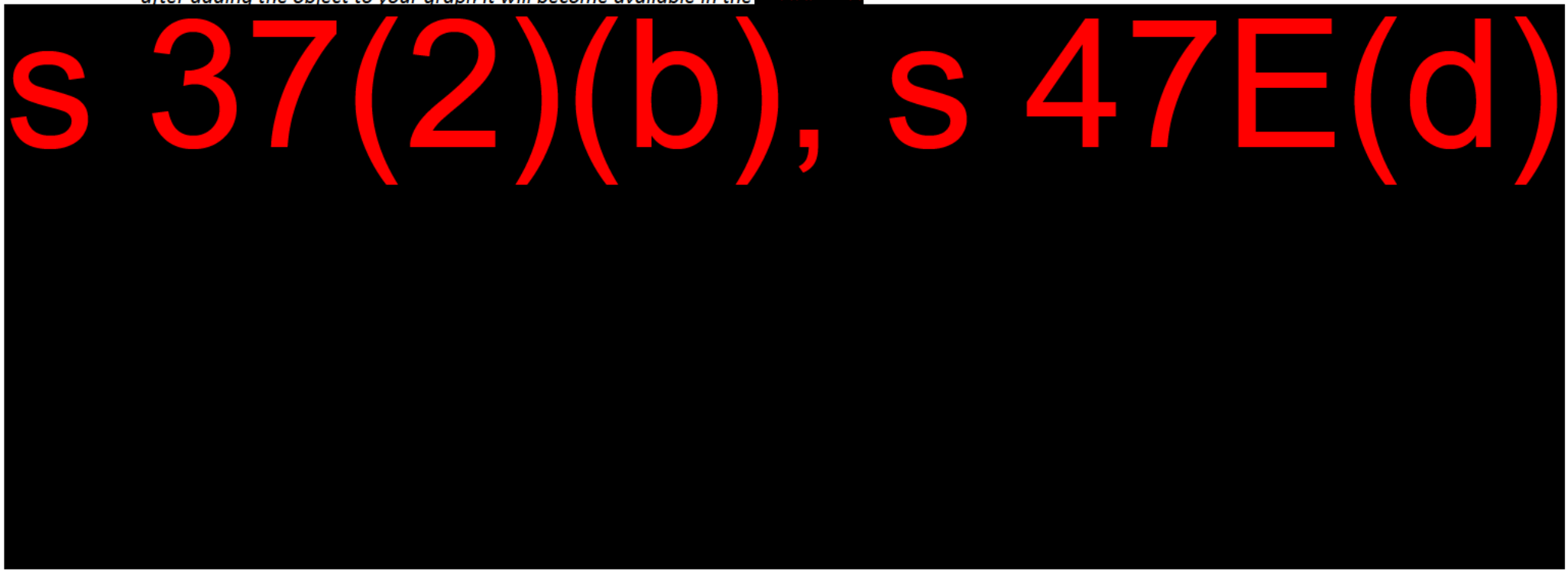
s 47E(d), s 37(2)(b) [redacted]

s 47E(d) [redacted] As you can see, there are various s 47E(d) templates to the left-hand side of the screen that can narrow down searches to specific entities. For this example, we will complete a general search using the s 47E(d) General Search' template

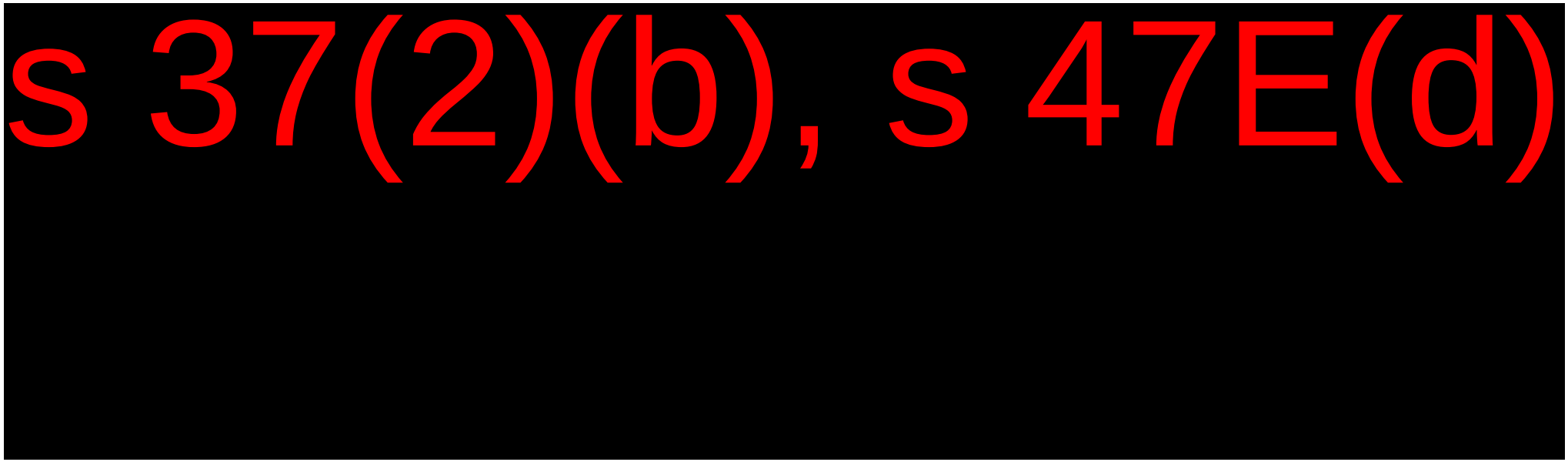
a. Enter the search criteria e.g. Bart Simpson
b. Enter a date range (if applicable)
c. You can also select the various data sources available to you from the **'Data source' drop down** (leave this as the default to search all datasets)
   e.g. only search s 37(2)(b), s 47E(d) [redacted]

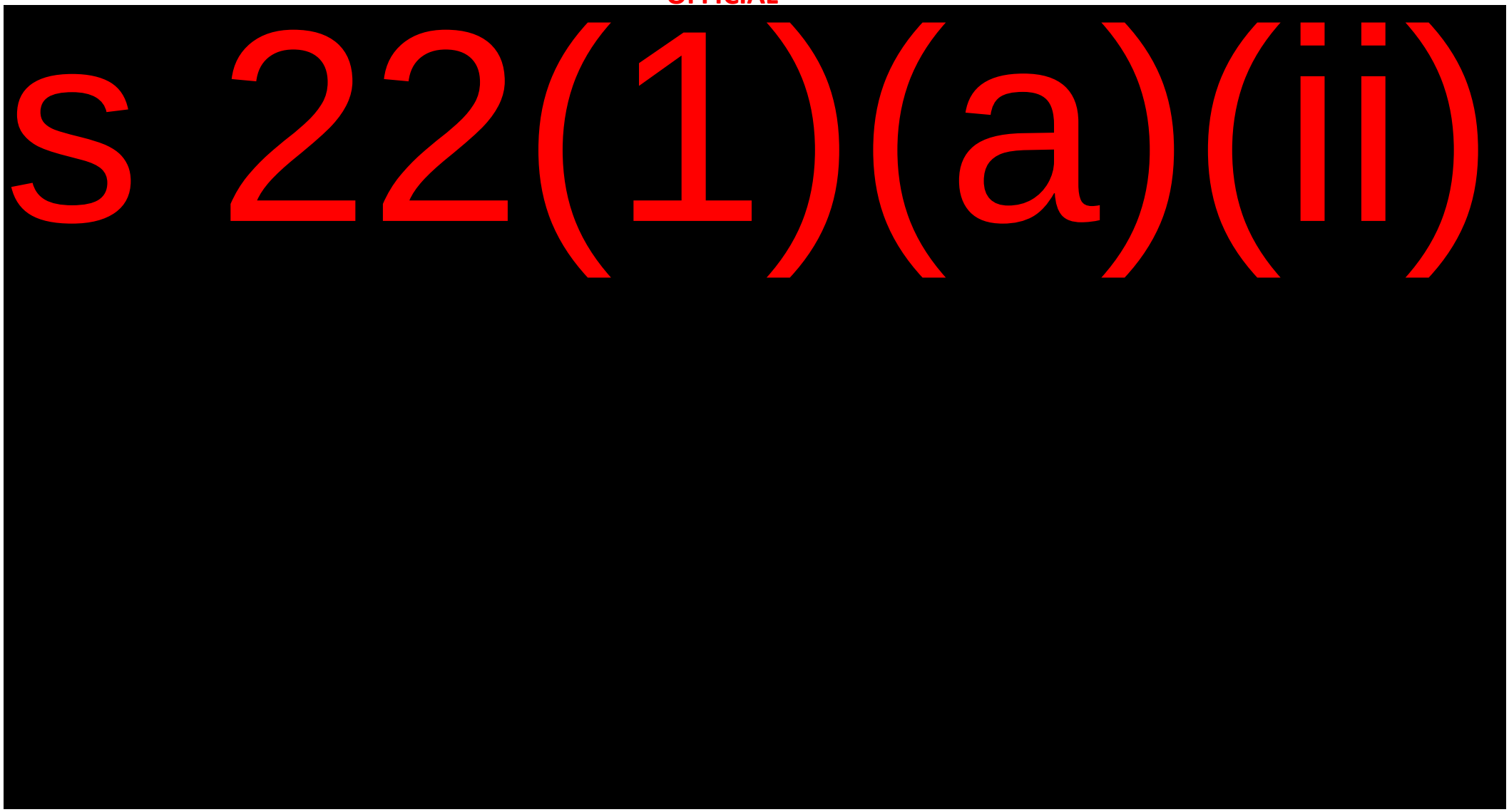


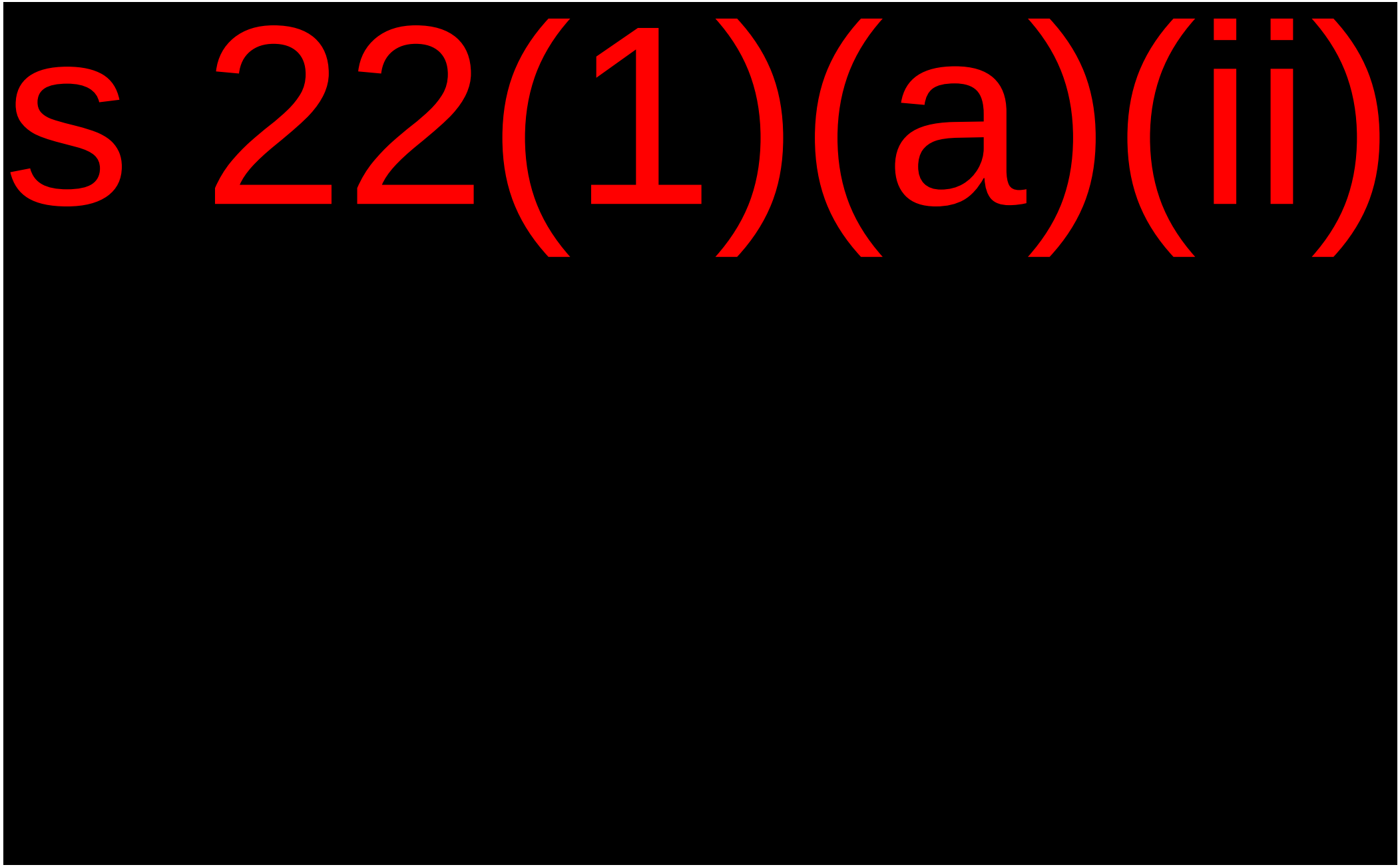
4. Your results will display below the search menu in order of relevancy **(represented by a score out of 10).** This will also be colour-coded, with 'green' indicating the highest score
    a. After you select your result, you can right-click to see a Preview of that information to the right-hand side of the screen *(this is particularly useful to **check the date of the source**)*
    b. You can right-click to add that object to your Graph application to add this to your chart for further analysis *(if you publish your investigation after adding the object to your graph it will become available in the* s 37(2)(b), s 47E(d)

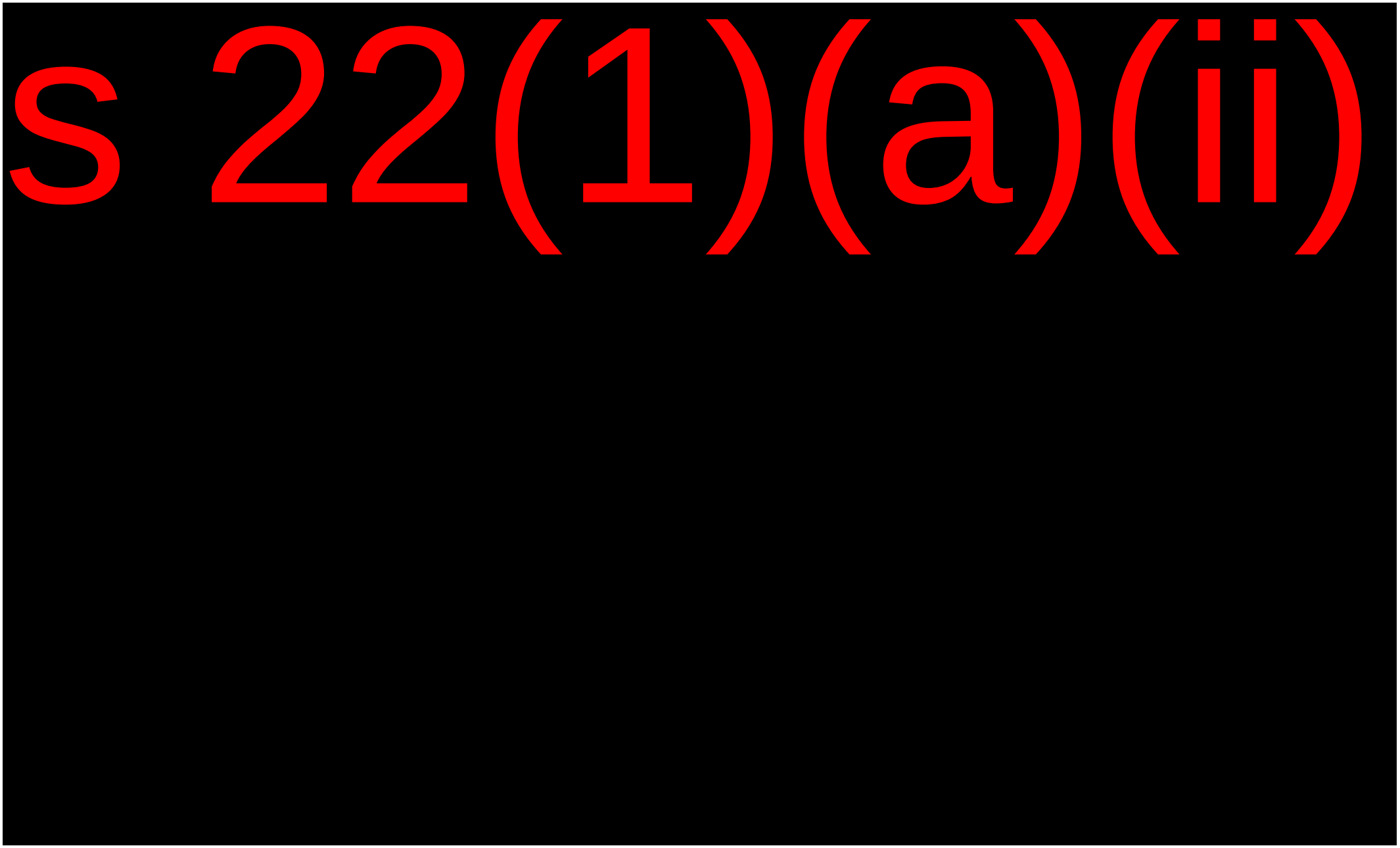


5. You can also use the filters in s 47E(d), s 37(2)(b) to refine your results. Simply click on the column header you wish to apply the filter to *e.g. the Name column*, using the drop-down arrow and select your desired filter *(results will still be displayed in order of relevancy after applying a filter)*




s 22(1)(a)(ii)


s 22(1)(a)(ii)


s 22(1)(a)(ii)

Step 3: Advanced Searches in Palantir

You can use advanced search operators in Palantir (including within the s 47E(d), s 37(2)(b)). These search operators can be used to narrow and expand your search results when dealing with large data sets and common words for example. Search operators can be used in a simple search and in the s 47E(d), s 37(2) application.

For the purpose of this user guide we will cover *four* common search operators to assist your searches in Palantir. If you are interested in learning more about these useful searching techniques outside of this guide, we encourage you to Google 'advanced search operators/syntax' to see further examples.

Boolean Operators

You can use the words AND, OR and NOT between keywords to both narrow and expand your results. **AND** will narrow your search by including *all* key words used. **OR** will expand your search so all returned results must contain at least one of your key words (helpful when searching for synonyms). **NOT** will narrow your search by eliminating all key words that follow.

For example – quick fox AND red fox

You can also set Parentheses when using Boolean operators to set the order of your search. Parentheses should generally be used around OR statements so your search can attempt to resolve the OR operator first, before moving to the next key word.

For example – (quick OR red) AND fox

Exact Phrases

You can use double quotes at the beginning and end of your search term(s) to search for that exact sequence in a document.

For example – "quick red fox"

The above example would only return results with "quick red fox" exactly

Wildcards

You can use '**?**' to replace a single character, or ***** to replace zero or more characters. The **question mark** character is best used when there are variable spellings for a word

For example – colo?r will produce results with 'colour' and 'color'.

The **asterisk character** is best used following the end of a root word (known as a 'truncation')

For example – gam* will produce results such as 'game' and 'gaming'

Additionally, both of these wildcards can be used in the same key word search to expand your results

For example – qu?ck* will produce results with 'quack', 'quacks', 'quick', 'quickly' etc.

Fuzzy Operators

You can use **~** (tilt character) before a simple search term to search for phonetically similar terms in addition to exact phrases. This can be especially useful when dealing with lengthy words where there's potential for misspelling, as well as names that have multiple spellings but sound alike *e.g. Mohammed vs Muhammed*

For example – ~quikc ~rd ~foks

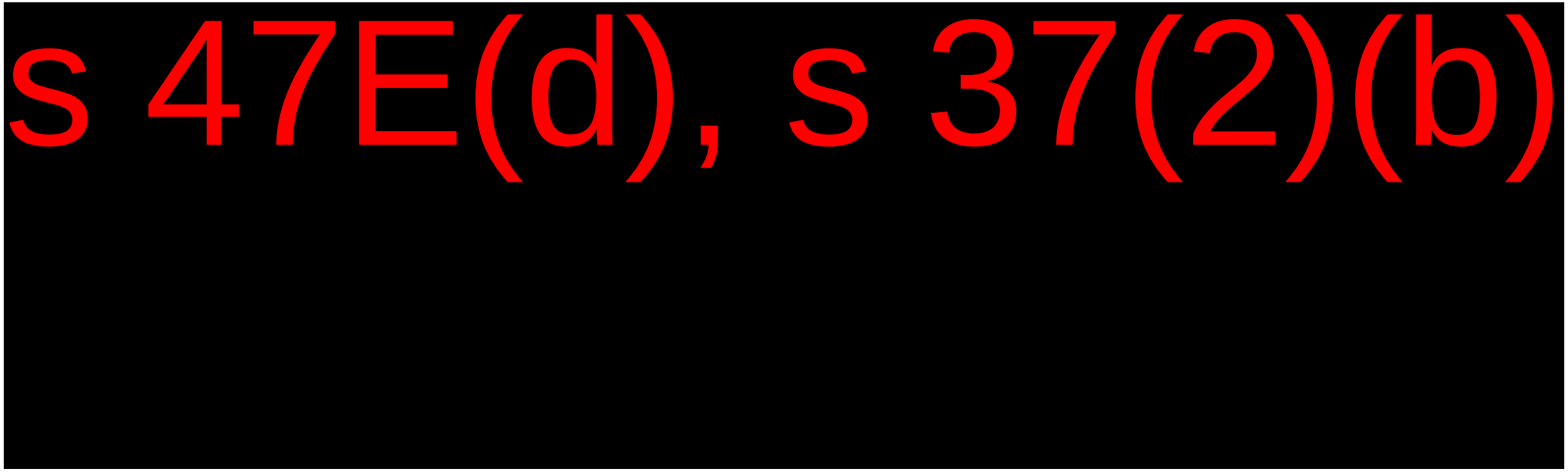
This search would return results with 'quick', 'red', and 'fox'

Please Note: You can also use the wildcard operator '*****' at the beginning of your search

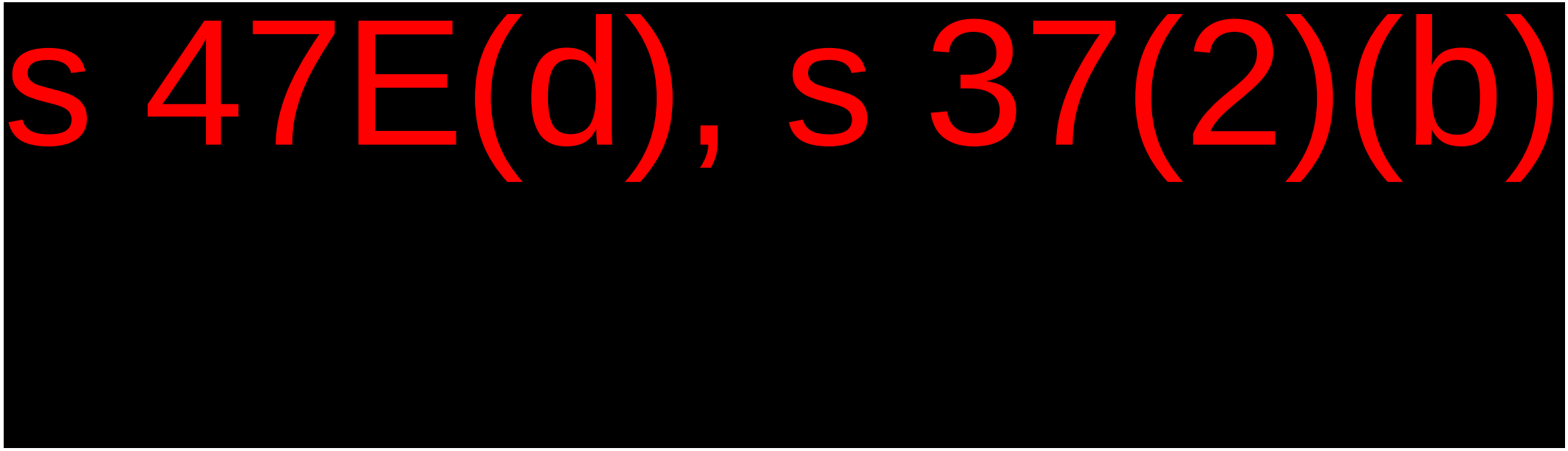
Step 4: Search Around Tool

The **Search Around tool** is only available in the Graph application and allows you to conduct search on a selected object using the Palantir s 47E(d), s 37(2)(b) to seek other linked objects such as entities, documents, events. You can also search for entities that have common properties with your current object.

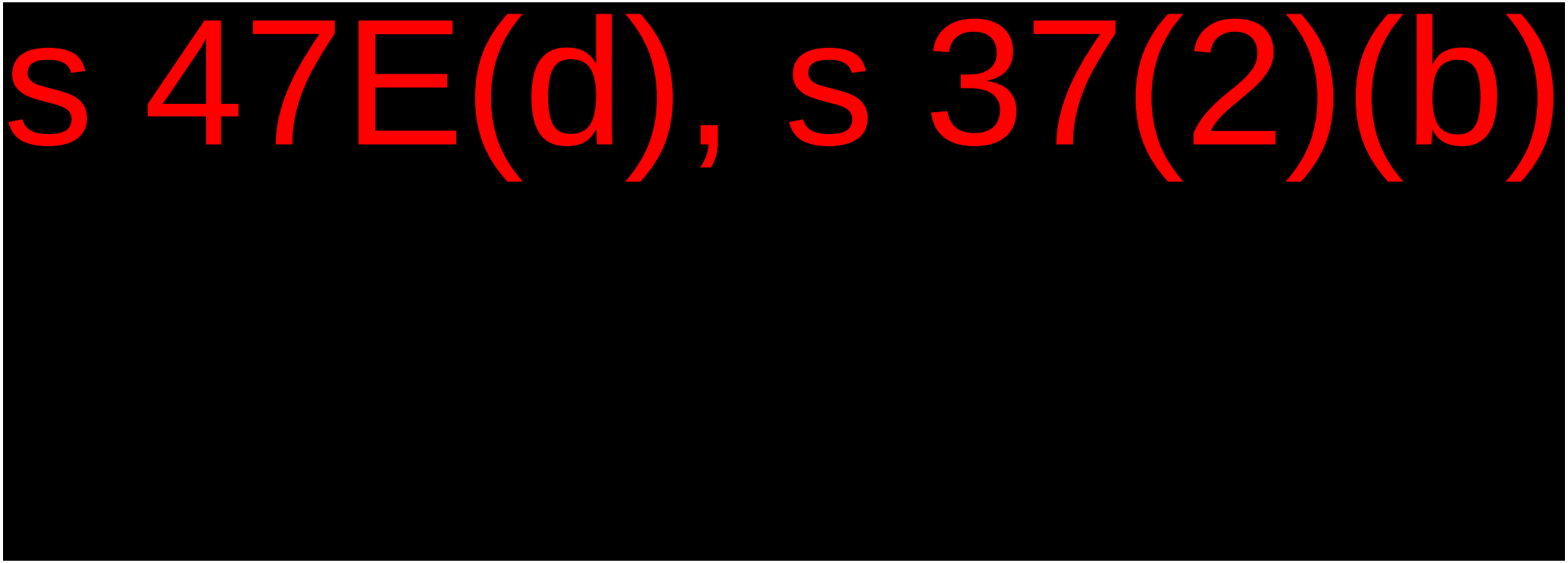
1. Using the **Graph application,** select your object and right-click to bring up the 'Search Around' tool



2. **Left-click** the **'search around' button** and select from your various actions *e.g. 'Linked entities'*
    1. The other search around templates include 'Linked Entities' and 'Linked Documents' and function the same as the 'Linked Events' used in this instance



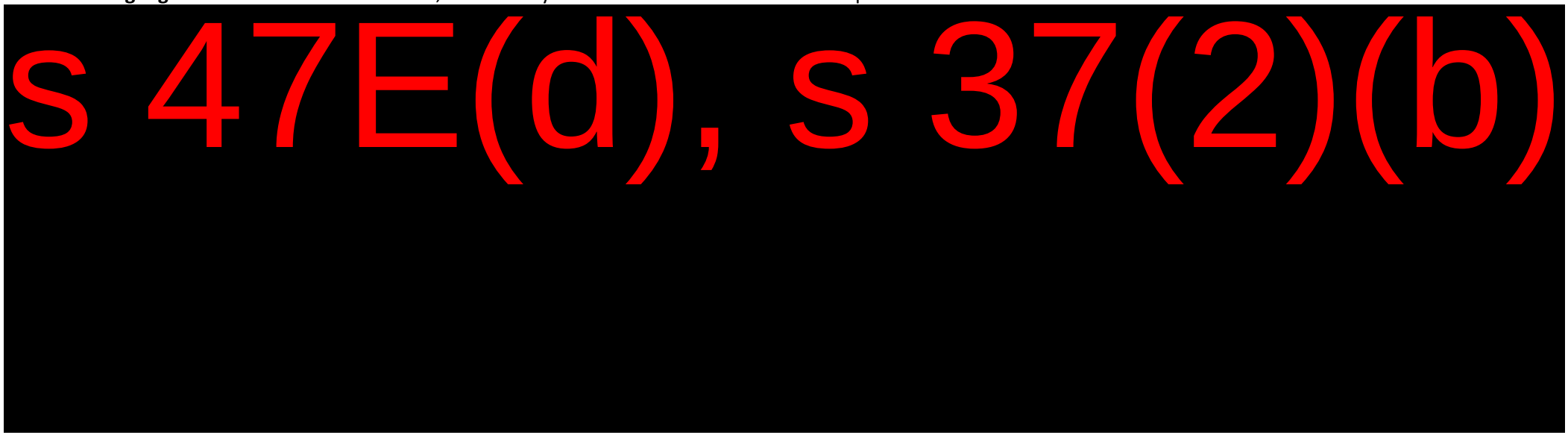
3. Once the search around is complete, it will now load *all* linked entities to your current graph for further analysis



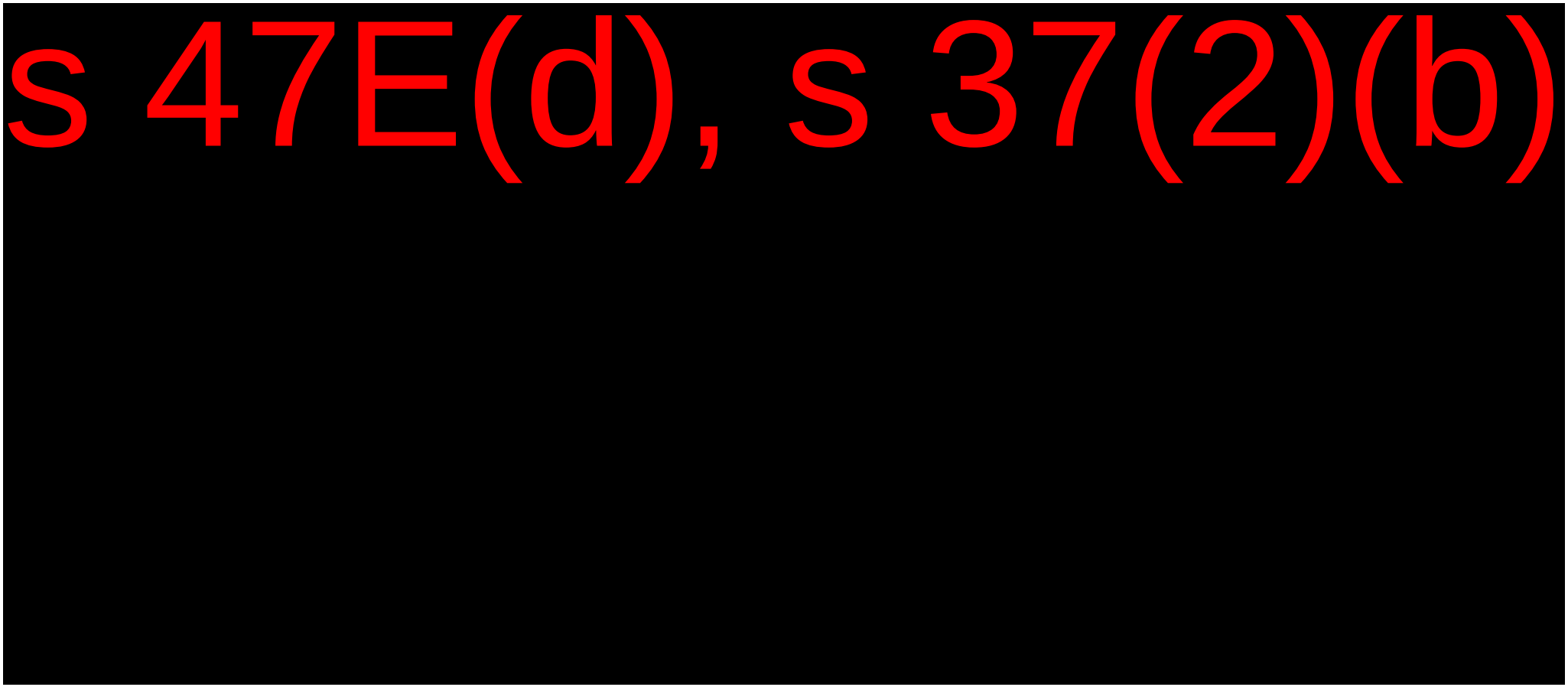
Step 5: Customised Searches in Palantir

You can complete a **customised search** in Palantir while in the Graph application. This uses the **'Search Around' tool** as seen in step 5 above, but allows you to modify your search terms, rather than use a pre-existing search around template *e.g. linked entities, linked events and linked documents*

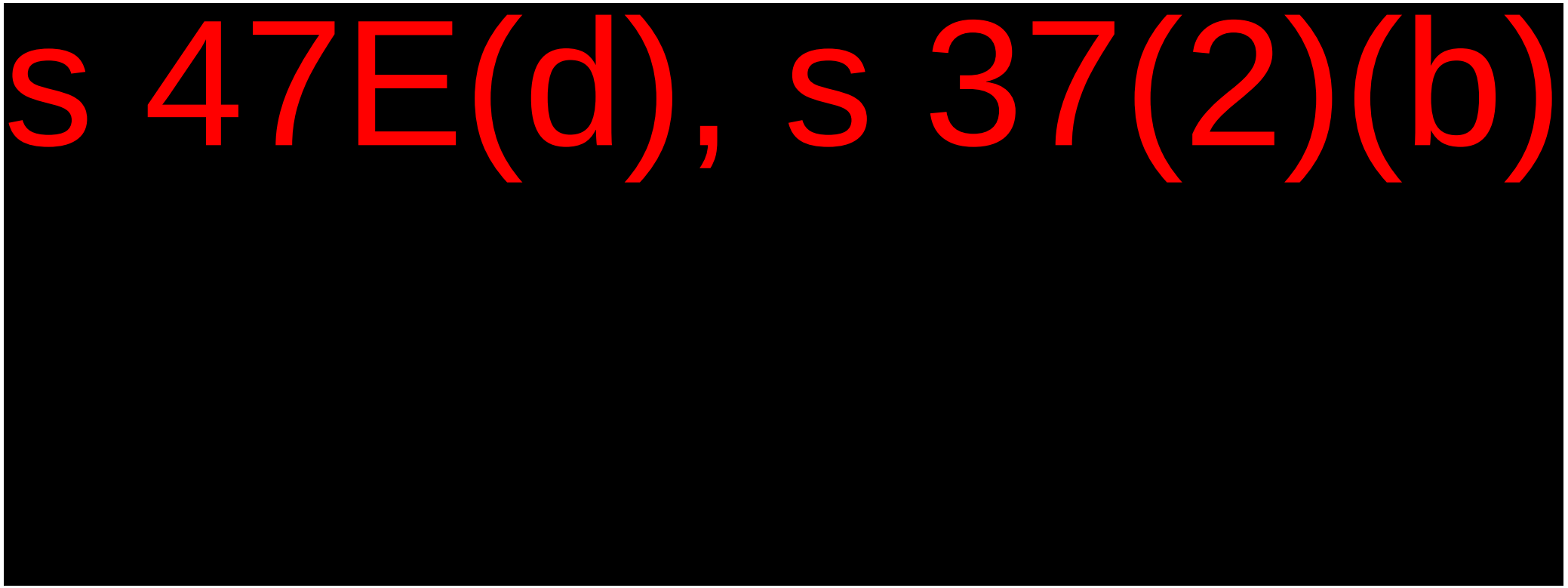
1. **Select an object(s)** on your Graph and **right-click** to bring up the search around tool
2. **Highlight the 'search around' button**, followed by the **'more…' button** and click to open a new window



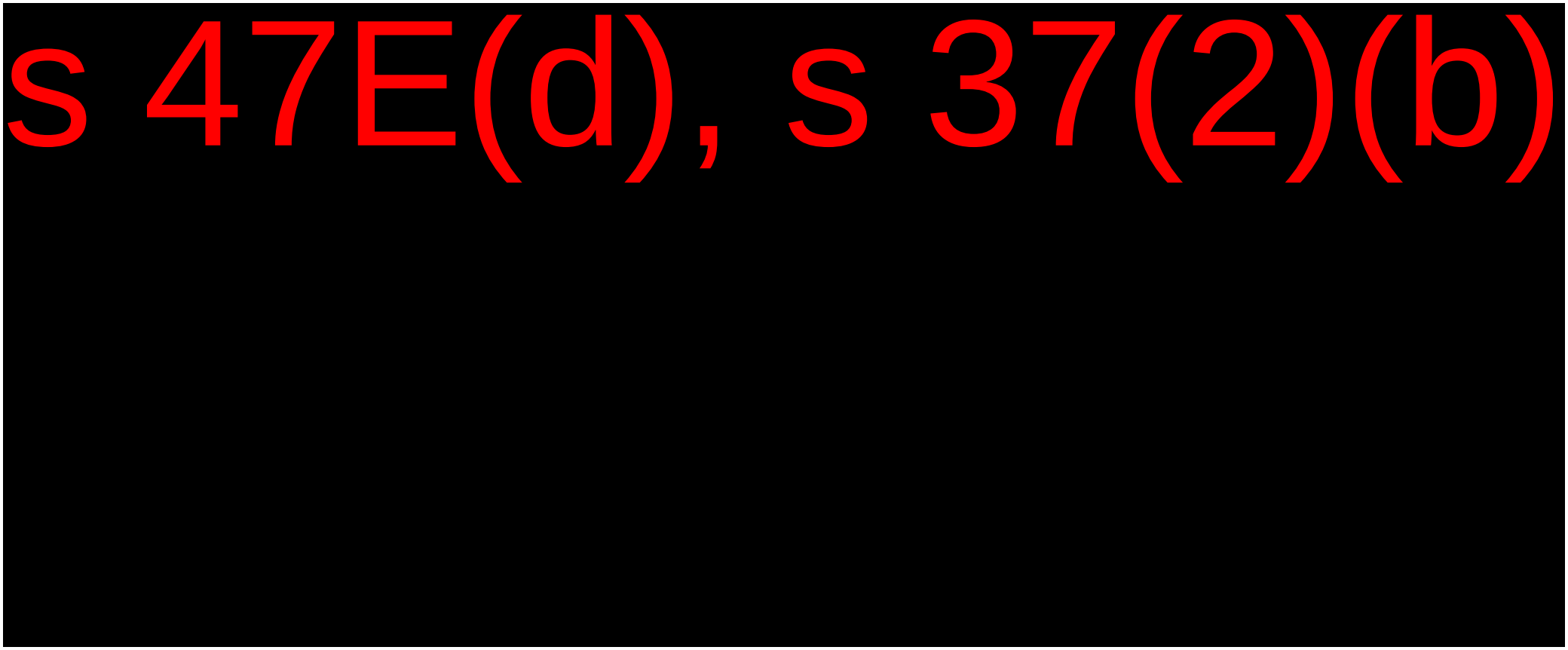
3. You can now select from numerous search templates. We will select **'Create a New Search'** in this instance



4. You can now select the objects you wish to create a search for. In this instance we only have one object on the graph so it has been selected automatically

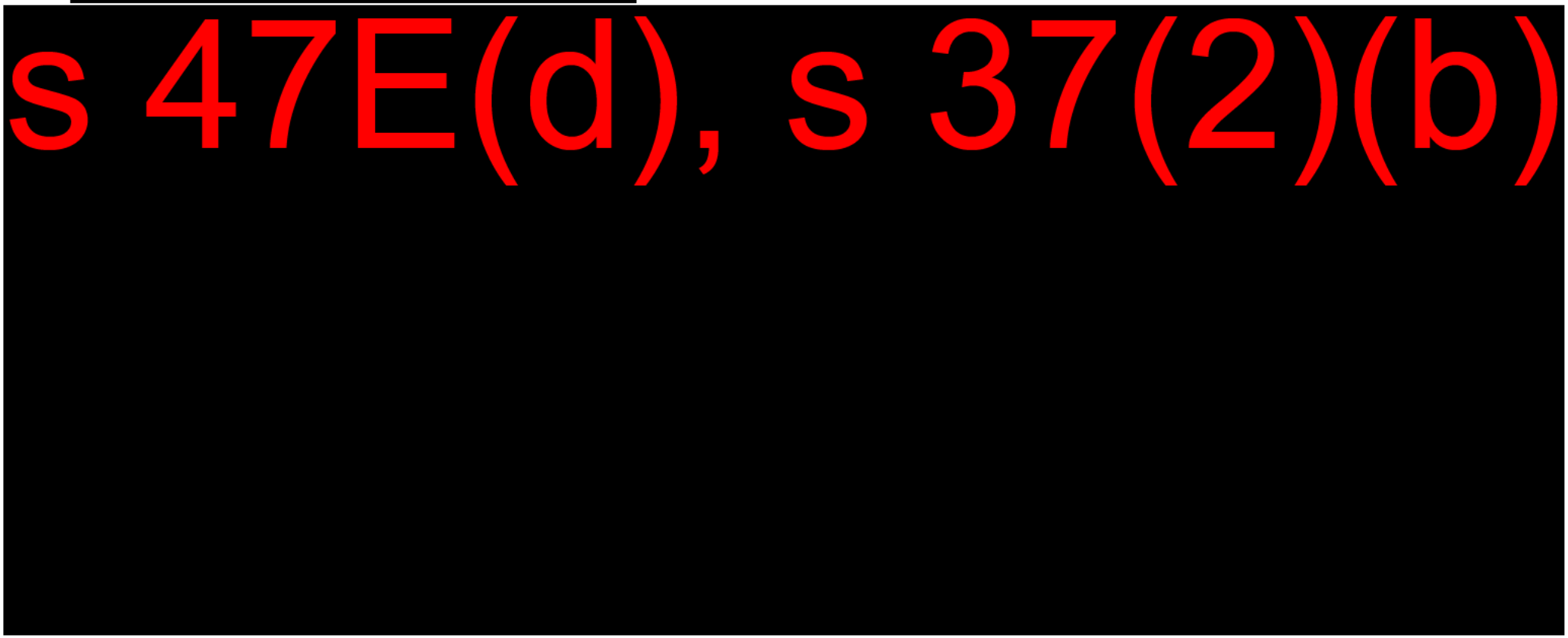


5. You can now select what link types and properties you wish to search for that may be connected to your object. s 37(2)(b), s 47E(d)

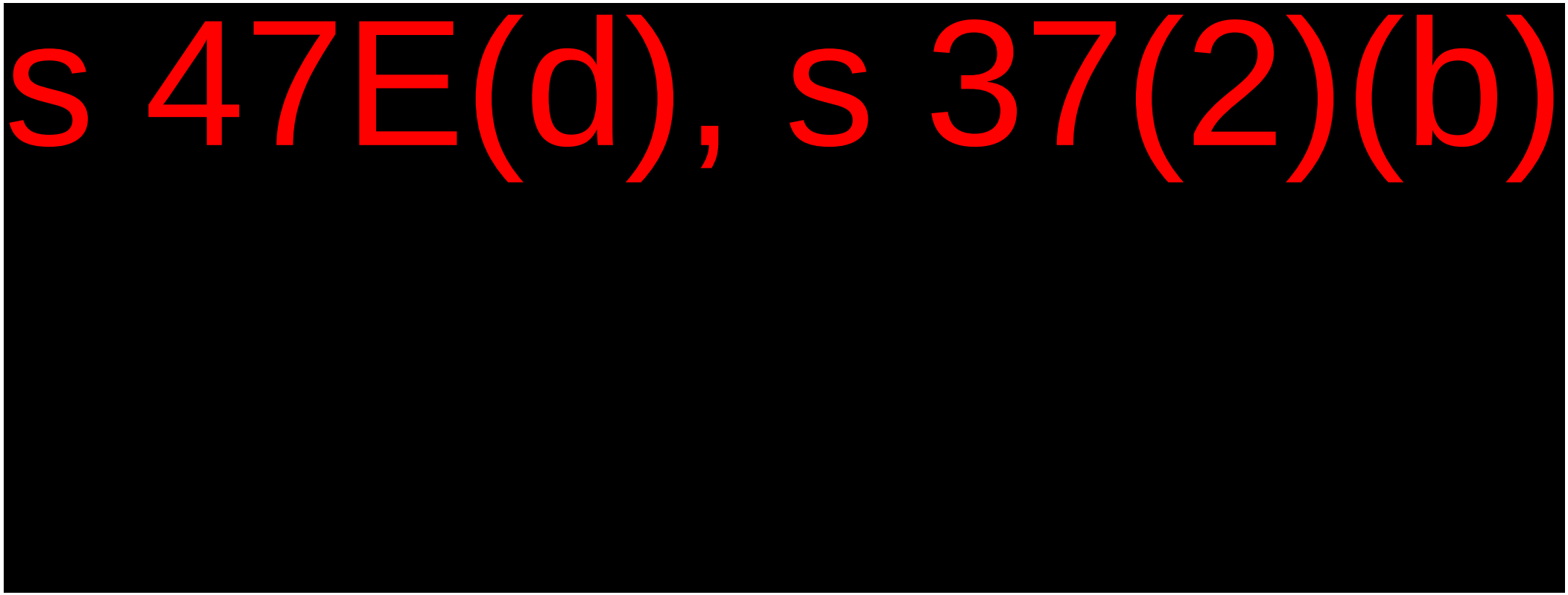


6. You can now select the target object types you wish to search for within your link type parameters. In [redacted]

[redacted]



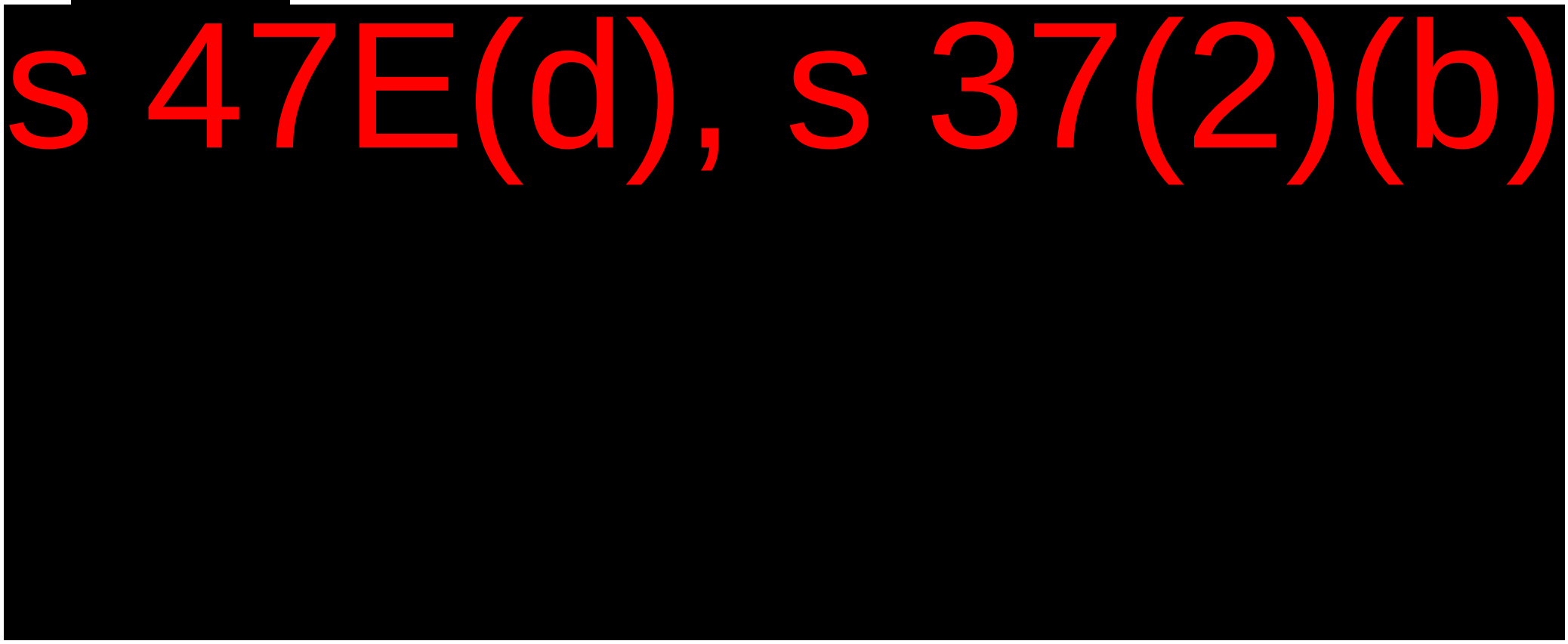
7. You will now see a **preview of your search results**. As can be seen below the customised search provided two linked entities within our search parameters s 37(2)(b), s 47E(d)
    a. To **save this customised search** as a template, click on the **'Save Copy as New Search'** which will appear at the top of your list next time
    b. **Click 'Finish' to complete** the search and load the new objects and links into your graph



Step 6: Object Explorer

The **Object Explorer** is an application that allows a user to search for *all* objects published to the Palantir [redacted] and apply filters to these objects to drilldown on specific entity or property types. This is similar to using the Histogram helper but opens your search results to every single object.

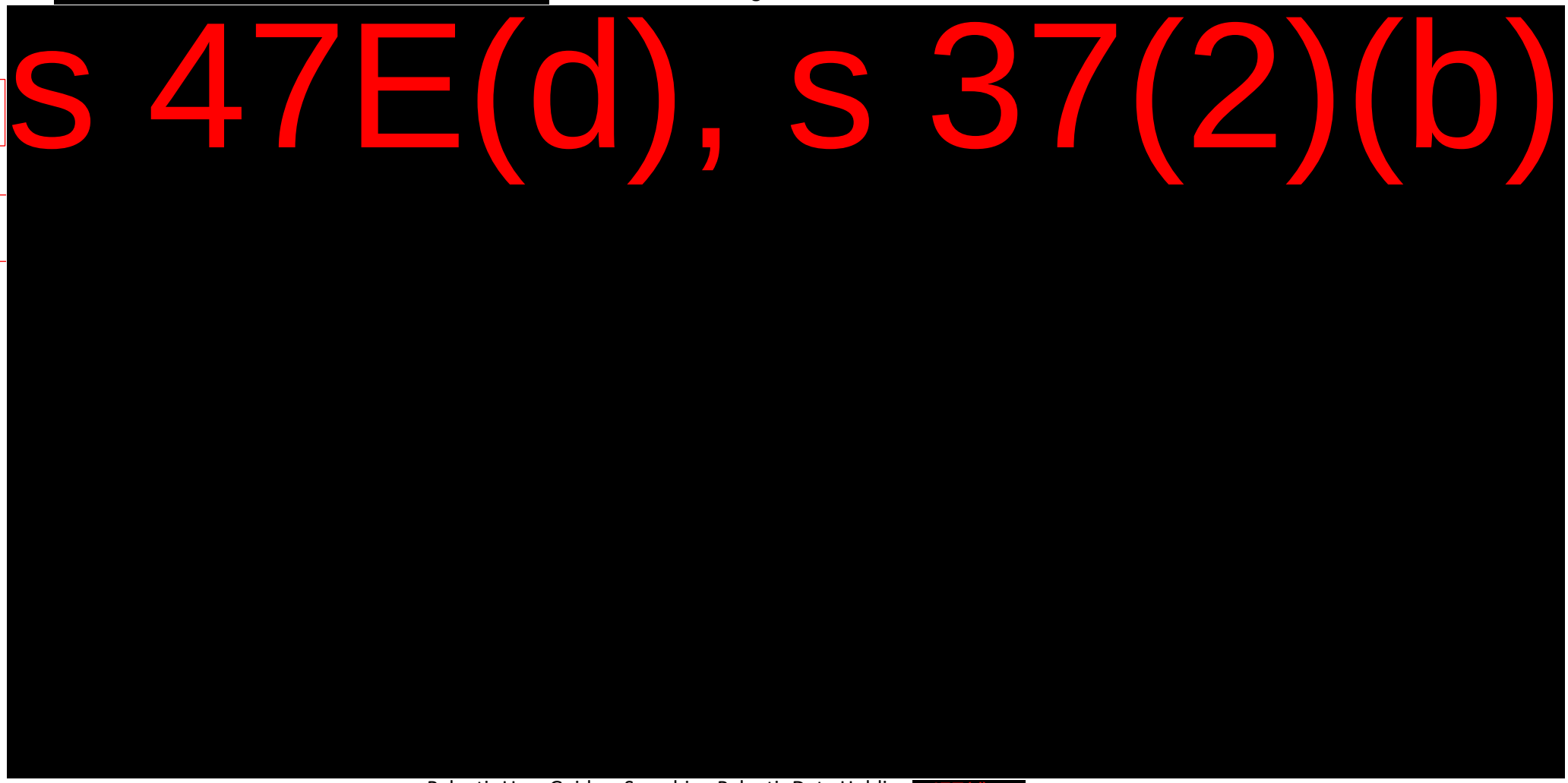
1. Activate the Object Explorer via the **'Application' toolbar**
2. When you open the Object Explorer for the first time you will need to click '**Explore All Objects**'. T[redacted]



3. Using the histogram-like **'type model' to the right-hand side of screen**, you can select objects and property types to search (you can also drag objects from other applications into the Object Explorer)
4. As you build your search you can track your stages to the top most part of the screen. s 37(2)(b), s 47E(d) [redacted]
5. You can apply various **visualisations** or **filters** to your searches as well. These can include charts, properties, timelines etc. s 37(2)(b), s 47E(d) [redacted] This visualisation will generate in the centre of the screen

4

5



End of Document

Last Modified: 28 Apr 2022

s 47E(d), s 37(2)(b) User's Manual

Introduction 2
Startup 2
Graphical Interface Components 4
Templates Panel 4
Search Panel 5
Results Panel 6
Preview Panel 8
Change History 8

Introduction

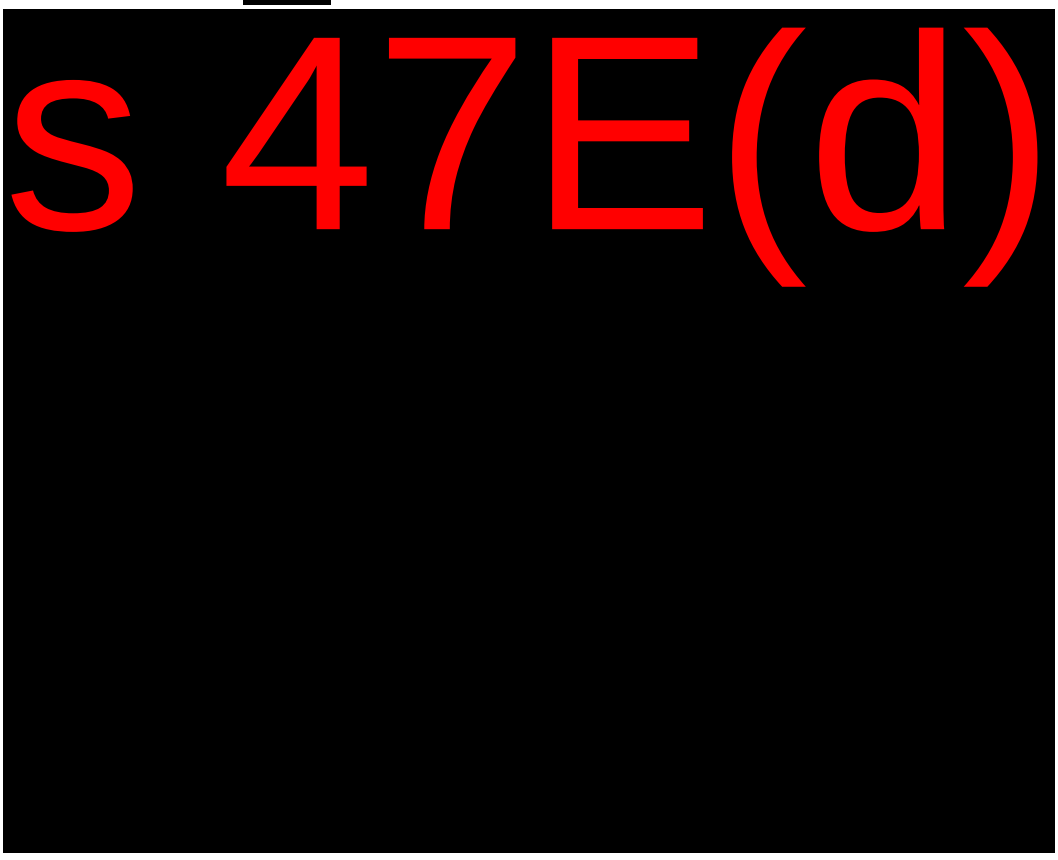
s 47E(d), s 37(2)(b) that enables users to search data in s 47E(d) Search results containing data from the s 47E(d) can then be displayed in Workspace Graph as entities and links, assuming a mapping has been created using Semantic Mapper.

Startup

In order to start Palantir Workspace with s 47E(d), click the magnifying glass in the lower left corner next to the Windows button:



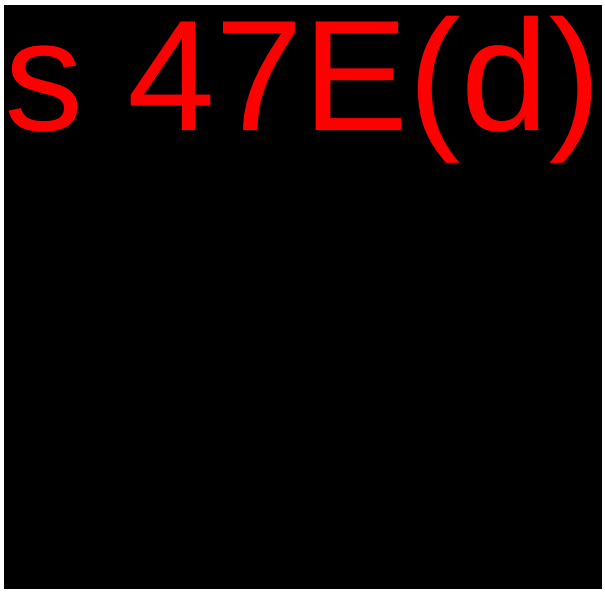
Click the Palantir s 37(2)(b), s 47E(d) or Palantir icon:



Click the Run button in the Open File popup window:



Palantir Workspace client starts up and displays the login form:



Once Workspace displays on the screen, check its toolbar to see if s 47E(d), s 37(2)(b) is visible (a magnifying glass):



and if not, click the Applications option on the menu and then select s 47E(d), s 37(2)(b) option:



Finally, s 47E(d), s 37(2)(b) is displayed within Palantir Workspace:



Graphical Interface Components

The s 47E(d), s 37(2)(b) GUI consists of the Templates panel on the left, the Search panel in centre-north, the Results panel in centre-south below the Search panel, and the Preview panel on the right (Fig. 6).

Templates Panel

The Templates panel contains user-selectable search types: document search (General, Person, Address, Organisation, ABN-CAN, Phone Number, Email) s 22(1)(a)(ii) .

Document Search is based on the index created by a [redacted] search engine. The user specifies keywords [redacted] finds [redacted] index documents containing those keywords. The [redacted] search engine creates one index document for every row in each document uploaded to the FDH ([redacted]). Each [redacted] index document contains fields with names and values. [redacted] finds the keywords specified by the user in the field values, and displays the whole index document for every match found. The latest version of [redacted] has a search result pagination implemented for all Document Search type searches [redacted] Users can define the search results page size, which persists across [redacted] sessions.

[redacted]

Search Panel

The Search panel contains graphical components for the user to specify search criteria and run a search.

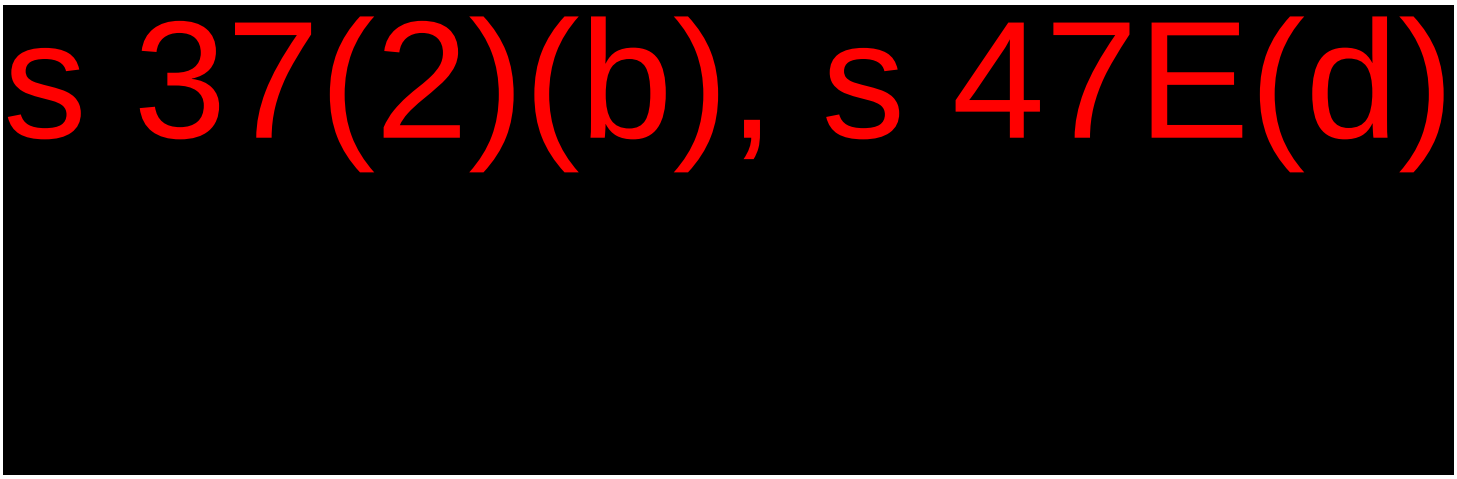
[redacted]

Those components depend on a search type. For instance, the [redacted] General search (Fig. 7) contains the Search field where the user types in one or more search keywords, the Upload Date group of radio buttons to specify an index-uploaded date range for search results, and the Data source pull-down list for specifying where the searched data comes from.

The Upload Date component is common for all Document searches. It allows the user to specify a date when each result was added to the index. The index upload date has been chosen because it is readily available in the current system, but it can be changed if required.

The Data source field is also common for all search types.

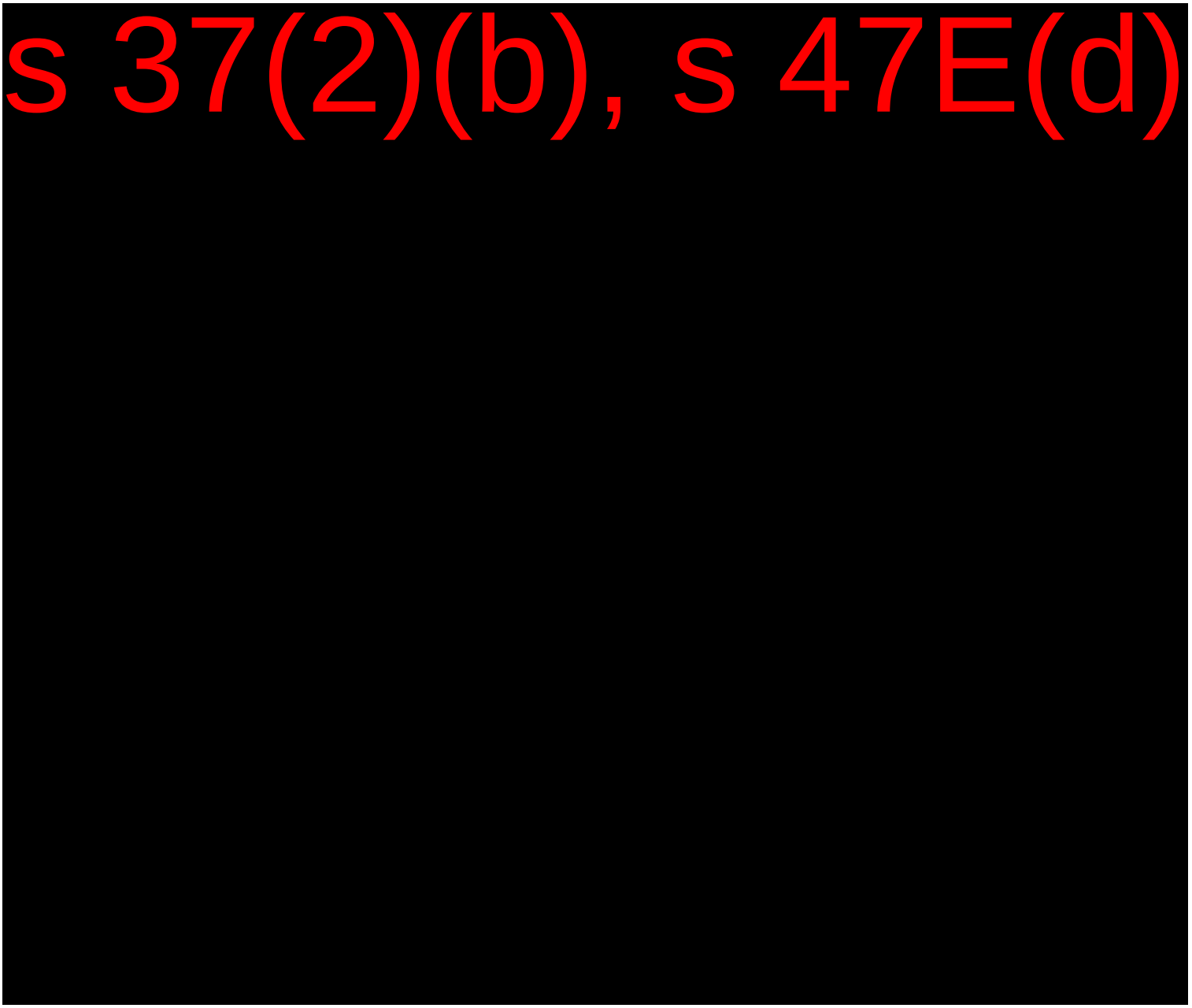
The Search panel contains the Status panel at the bottom, where the total count of search results will be displayed for all search types, including the search time:



In addition, the Status panel contains result pagination components for Document searches. The result page size is user-defined and it defaults to 35 rows initially. s 37(2)(b), s 47E(d)

Results Panel

The Results panel displays search results sorted by their score in the descending order:



Search result rows are colour-coded. The results with the score equal to the maximum score (for this current search), are in green. The results with the score in the range (75%, 100%) (left inclusive, right exclusive) are in yellow. The results with the score in the range (67%, 75%) are in light grey. The rest of the results are in dark grey colour.

Double click a particular result row in order to display a preview of the selected document. Right click it and then select the Add to Graph option to add the selected document to Workspace's Graph.

Preview Panel

The Preview panel (Fig. 10) displays user-defined tags for the selected document (s 37(2)(b), s 47E(d) and document's row of data with column names.

s 37(2)(b), s 47E(d)

Change History

Version 11.0.3 s 47E(d)

Compacted some GUI elements (date panel, button panel) to occupy less vertical space in SearchPanel to address the issue of the results panel being too small for lower screen resolutions or larger font sizes.

Version 11.0.2 s 47E(d)

s 22(1)(a)(ii)

Version 9.7.4 s 47E(d)

s 22(1)(a)(ii)

Version 9.7.3 s 47E(d)

s 22(1)(a)(ii)

s 22(1)(a)(ii)

Version 9.7.2 s 47E(d)

s 22(1)(a)(ii)

Version 9.7.1 s 47E(d)

s 22(1)(a)(ii)

Version 9.7.0 s 47E(d)

Removal of s 47E(d)

Version 9.6.1 s 47E(d)

a) Restored resizing of the preview panel
b) Restored full Lucene query syntax (with double quotes, etc.)
c) Implemented user-defined search result page size
d) Improved preview panel population

Version 9.6.0 (3 Apr 2017)

a) Addition of the Upload Date field for Document searches
b) Search result pagination for Document searches
c) The result status panel shows the total number of search results
d) Removed glass pane (temporary greyed out s 47E(d), s 37(2)(b) panel while searching)
e) Replaced 'NULL' strings with empty strings in the search result panel
f) The Search Criteria panel resizes dynamically depending on the selected search template
g) Score values are formatted with only one fraction digit
h) Increased the Identity search timeout from 15 to 30 seconds

AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

BULK DATA CLEANING GUIDE

Bulk Data Collection

MAY 2026

CONTENTS

Data Cleaning Guide **3**
Important Notes Before You Start 3
Changing Delimiter 3
Adding a SORT Column 3
Removing Multiple Spaces 4
General Use 4
Tools for Identifying Errors 4
Formatting Tools 4
Converting Tabular Data from PDF to Excel **6**
Entity Cleaning **7**
Individual & Organisation Names 7
Split Organisation from Individual Names 7
Find Organisation Names Split between Two Columns 8
Identify Names with 'Last, First' Structure 8
Identify Semi-colons in Name Cells 9
Phone Number 10
Reformat Australian Mobile Numbers 11
Reformat Scientific Notation 13
Dates 13
Reformat Dates with a Timestamp 14
Reformat YYYY-MM-DD 16
Addresses 17
Postcodes 18
Emails/Websites 18
Seperating/Concatenating Data **20**
Separating Data 20
Separating Phone Numbers in Free Text Field 20
Combining Data 22
Merging Cells 22
Merging Columns 22
Saving Files **24**

Formatting CSV Files 24
Saving Files in CM 24
Naming Conventions 24
Keyboard Shortcuts **25**
Quality Assurance **26**
QA Checklist 26
Tracker 27
Appendix **28**
Appendix A – Excel Functions 28
Appendix B – Irrelevant Data 30
Appendix C – Phone Number List Separators 30
Appendix D – Keywords used to identify Company Names 30
Appendix E – Keywords used to identify Multiple Names 2
Version Control **2**

DATA CLEANING GUIDE

This document contains procedures for cleaning bulk datasets to ready them for ingestion into the [redacted] database.

Our goal is to upload data of the highest quality for the benefit of intelligence analysts. It is therefore important that throughout the cleaning process we maintain the integrity of the data. We do not alter records themselves but ensure that data is accurately categorised, properly formatted and free of null values.

This guide addresses the cleaning of fields/categories that occur frequently and have a lengthy and repeated cleaning process. Note this guide is not exhaustive - google is great for finding new ways to clean data or build on the functions mentioned below.

It is important to balance cleaning efforts with the constraints of time. Whist this guide outlines various methods for identifying erroneous data it is acknowledged that some errors will be overlooked in any given dataset. This is mitigated by a comprehensive review process outlined in the QA section below.

Note: All examples of data in this document are purely fictional.

Important Notes Before You Start

Changing Delimiter

[redacted] uses Pipes as a delimiter. Make sure your list separator has been changed from comma "," to pipe "|". To do this:

1. Date and time settings
2. Language & region
3. Administrative language settings
4. Format
5. Additional settings
6. List separator:
7. Change this from "," to "|"

Adding a SORT Column

The first thing to do when cleaning bulk data is to insert a SORT column with ascending dates (from 01/01/2000) at column A. This helps preserve the original order of the data as it is received. After sorting or filtering, rows can be reordered (sort by A-Z) for easy comparison with the raw data.



Removing Multiple Spaces

Replace multiple spaces with single spaces by using the Find and Replace tool. Replace “ ” with “ ”. Repeat this process until Excel can no longer find anything to replace. Note sometimes Excel will continue to display the same message (e.g We made XX replacements) indicating this process is done.

General Use

Cleaning requires detecting problematic data and then transforming this in some way. Below is a list of frequently used tools and functions that can be used to easily identify errors.

Tools for Identifying Errors

Filters can help check columns for data that shouldn’t be in the cell (e.g., spaces in phone numbers, “&” in names). It is important to be careful when using filters as pasting and deleting can affect cells hidden by the filter.

Ctrl+A Sort brings problem cells to the top or bottom of the column by sorting A-Z. Numbers or special characters will be arranged above alphabetic characters. This quickly identifies numbers in a name column, names in a number column, etc.

Find and Replace (Ctrl+H). This is useful for changing commas to semicolons and removing full stops and other characters when necessary. Use with caution as this can delete unintended stuff (e.g., replacing “.” with nothing will get rid of “.” in phone numbers but it may also remove it from emails if you haven’t selected the right column).

Removing line breaks - While in the find text box enter Ctrl+J. Sometimes this does not work en masse. See unwrapping special characters in the section below.

=IF can be used to identify which cells needs to be cleaned by producing a true and false value, e.g., =IF(ISBLANK(*cell*|”X”|””).

=LEN can be used to measure the length of each cell. Extremely long or short cells are often an indicator of incorrect data. These outliers can be found by using filters.

Formatting Tools

Wrap Text button (found in the alignment section of the home tab) will unwrap text to a single line format. Occasionally data contains special character line breaks that cannot be unwrapped using this tool. In this instance:

1. Highlight the column
2. Select txt to column
3. Select Data
4. Select 1st option Delimited
5. Select Next
6. Choose Other
7. In the box next to other: Hold down the ALT Key and type 010
8. Select Next and Finish

=TRIM removes spaces at the beginning and end of words.

CONVERTING TABULAR DATA FROM PDF TO EXCEL

When data is provided in a table in PDF format, use the following instructions to convert it into Excel spreadsheets.

1. Save the PDF containing the data you want to your **local desktop**. (**Important that it is saved to the local desktop**).

2. Open a new Excel spreadsheet, go the "**Data**" tab and select the "**Get Data**" icon

3. After selecting the "**Get Data**" icon, click on the "**From File**" drop down

4. Within the "**From File**" drop down, click on the "**From PDF**" option

5. Select the desired PDF on your local desktop and wait for a connection to establish between the Excel and PDF documents

6. Scroll through the window containing the recognised tables within the PDF and select the desired option. Avoid selecting multiple items at the same time, it will cause collation issues.

7. Click the "**Load**" button and the table data will be copied and deposited into the Excel spreadsheet

8. If multiple tables of data are required from the same the PDF, open a new sheet within Excel and repeat the previous steps

For harder and more complicated datasets, please continue to send them to Advanced Analytics for processing and assistance.

ENTITY CLEANING

Individual & Organisation Names

Cleaning name columns is required in almost all datasets. Typically, the process consists of removing or relocating non-name data as well as splitting multiple names or corporation names into new columns.

Moving data that shouldn't be in the name column can be done by manually inspecting each cell (if the dataset is small enough) or using the processes listed below.

To identifying multiple names use Find (Ctrl+F), a filter or IF function to search for " and " (note that the spaces are important to avoid finding "Andrew") and other symbols listed in Appendix E. Businesses can also be identified and moved into a separate column by searching or filtering for "Pty.", "Ltd.", "Co." or variations thereof listed in Appendix D.



It is best practice where possible to separate multiple individuals (and any data associated with them) into new and corresponding columns. Where this is not feasible, a semicolon in between the names is acceptable.



Split Organisation from Individual Names

1. Review column for non-name data.
    a. Sort column by A-Z. This brings non-alphanumeric, followed by numeric characters to the top.
    b. Filter columns for 'Contains' = "@".
    c. Run an alpha check in a new column using a SEARCH function (Appendix A). Filter column for "FALSE" values to show cells containing numeric characters.
2. Delete irrelevant data (Appendix B) and relocate other data into appropriate column or a new column for further review.
3. Where the data does not inform you of the company/individual status; review for corporate entities and other organisation names.

a. Use a SEARCH function (Appendix A) in a new column that looks for common corporate keywords (Appendix D). Sort this column Z-A so TRUE values appear up the top.

4. Where cells are TRUE, cut and paste names into a new column.
    a. Review for errors.

Find Organisation Names Split between Two Columns

Used to find business names that have split been across a first and last name column.

1. Insert new column.
    a. Type the formula =IF(ISNUMBER(SEARCH(A1|B1))|"YES"|"NO").



    b. 'YES' or 'NO' will appear indicating whether there is a partial match between the two columns.
2. Double click in the corner of the cell to apply the formula to the entire column.
3. Select the column and sort by column C.
    a. This will sort all of the rows containing matches to the top of the column.



Identify Names with 'Last, First' Structure

1. Insert a new column.
    a. Type the formula =IFERROR(FIND(","|A2)|"none").



b. ‘None’ or a number indicating the position of the specific character will appear.

2. Double click in the corner of the cell to apply the formula to the entire column.
3. Select the column and sort by column B.
    a. This will sort all of the rows containing the specific character to the top of the column.



Identify Semi-colons in Name Cells

1. Insert new column.
2. Type the formula =LEN(A2)-LEN(SUBSTITUTE(A2|";"|"")).



    a. ‘0’ or a number indicating the amount of the specific character within the cell will appear.
    b. Double click in the corner of the cell to apply the formula to the entire column.
3. Select the spreadsheet and sort by column B.

a. This will sort all of the rows containing the specific character to the bottom of the column.



Phone Number

Phone numbers require cleaning to ensure uniformity where possible. Australian mobile numbers should be in a 10-digit format and usually require transformation (+61 format to be removed). Australian home phones are typically an 8-digit format. International phone numbers can vary in length – the smallest international phone number is 7-digits long whereas the largest can be up to 15. International phone numbers are left untouched however we do remove any "+" symbols.

Ideally, the phone number column would always be formatted as text to prevent Excel's automatic translation into scientific notation and automatic removal of leading zeros. Currently, there is no method to stop the conversion, particularly after any formatting or functions are applied to cells. Scientific notation that has been applied to data then saved as a CSV can led to data loss, as it will round to a few significant figures. You can tell this is the case where numbers have a long trail of zeros. These numbers need to be reviewed individually to determine whether a complete number is salvageable.

1. Review column for non-numeric data.
    a. Sort column by A-Z. This brings non-alphanumeric characters to the top.
    b. Sort column Z-A. This brings alphabet characters (starting at Z) to the top.
    c. Run a length check in a new column using the LEN function. Filter column for outliers or anomalies based on length (too short or too long).
2. Delete irrelevant data (Appendix B) and relocate other data into appropriate column or a new column for further review.
3. Review for phone numbers embedded amongst other data.
    a. Run a number check in a new column using a SEARCH function (Appendix A). Filter column for "FALSE" values to show cells containing non-numeric characters.
    b. Run a length check in a new column using the LEN function. Filter column for outliers or anomalies based on the length (too short or long). Very long numbers may be international numbers or may be random strings of numbers.
4. Ensure multiple phone numbers (occurring in a single cell) are split into a new column or are separated by ";" depending on the circumstances. Number columns should correspond to the

correct entity, e.g., It will be assumed that 'Phone Number 2' is linked to 'Contact 2'. Concatenate with a ";" where there is more than one number for one entity.

   a. For minimal occurrences - Manually separate numbers.

   b. For frequent or structured occurrences - Use Text-to-columns, a SEARCH function or the Find and Replace tool where a list has been separated by other characters (Appendix C), e.g., 'hm', 'wk'.

5. Move phone numbers identified in incorrect columns into the phone column. Note relocation becomes complicated by datasets with multiple phone number columns (primary or alternate).

   a. For minimal occurrences - Manually cut and paste two numbers adjoined by a semi-colon.

   b. For frequent or structured occurrences - Use a CONCAT function.

6. Re-format Australian mobile numbers to a 10-digit format (see instructions below).

7. Re-format cells in scientific format (see instructions below).

Reformat Australian Mobile Numbers

1. Select the column and change the format to 'Text'.

   a. Ignore any errors of 'Number stored as Text'.

2. Use the Find and Replace tool on the column to replace spaces with nothing. Also replace "-", "/", "+" and any other symbols (Appendix C) with nothing.

3. Sort the dataset from smallest to largest (or A to Z). Then copy and paste into a new sheet.

   a. Review column for non-numeric data. Remove and reallocate data where necessary. Sort again to move all the now empty rows.

4. In a new column enter the LEN function =(LEN(cell)). This will provide a count of the number of digits in each phone number.

   a. Double click in the corner of the cell to apply the formula to the entire column.

   b. Because the data was sorted before you copied it to this tab, the count of digits should be smallest to largest.

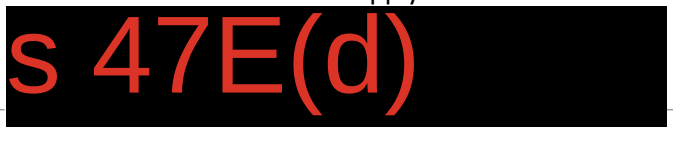
5. Scroll until you find cells with 9 digits that do not begin with a zero.



Example of =LEN in Column B

6. In a new column, enter the formula =BASE(phone number cell|10|10). This will put a zero at the beginning of the number.

   a. Double click in the corner of the cell to apply the formula to the entire column.



s 47E(d)

7. Copy and paste values over the old phone numbers to update them. There should now be 10 digits numbers beginning with zero.

   Easier Solution: Select all 9-digit numbers without '0', right click, select 'format cells', select 'custom', under 'type' type in '0#########'.

8. Scroll down until there are 11 digits in the cell and the cell begins with 61. Select all cells that apply.

9. Under the 'Data' menu, select 'Text to columns'.

   a. Select 'fixed width' and click 'next'.

   b. Click anywhere in the 'Data preview box' and adjust the line until it separates 61 and the remaining digits. Click 'next'.



   c. In the new pop-up box, for the first column selected under 'Column data format' select 'Do not import column (skip)'. In the same pop-up box, select the second column. Under 'Column data format' select 'Text'.

   d. Check the destination is column A (number of top selected cell), then click finish.

   e. This should remove the 61 from the selected phone numbers. **Note** you cannot use find and replace to do the same task because it would also remove any 61 in the phone numbers.

10. Repeat step 7 **'Easier solution'** to format these figures into 10-digit phone numbers beginning with zero.

11. Copy the column and paste it back into the first sheet. Delete the column.

Be Careful To:

- Check that there aren't multiple phone numbers in one cell and, if there is, that they are adequately separated with ";".
- Paste values when pasting the new numbers over the old numbers, otherwise you will lose these numbers.

- Ensure the format for the cells is “text”, otherwise you will lose the 0s on the front of the phone numbers.

Reformat Scientific Notation

1. Select the column, right-click and select ‘Format Cells’.
2. Under the ‘Number’ tab, decrease decimal places to zero.



3. Click ‘ok’ and the entire phone number will be displayed.



Dates

Dates are to be standardised to the dd/mm/yyyy format.

1. Select the column and change the format of the column to ‘Short Date’.
    a. Go to the ‘Number’ section of the ‘Home’ menu to find the drop-down box.
2. The format of the dates should convert to the correct format of dd/mm/yyyy, e.g., 01/01/2000.

There may be times where the above method does not work and dates do not convert to the correct format. See further instructions below.

Reformat Dates with a Timestamp



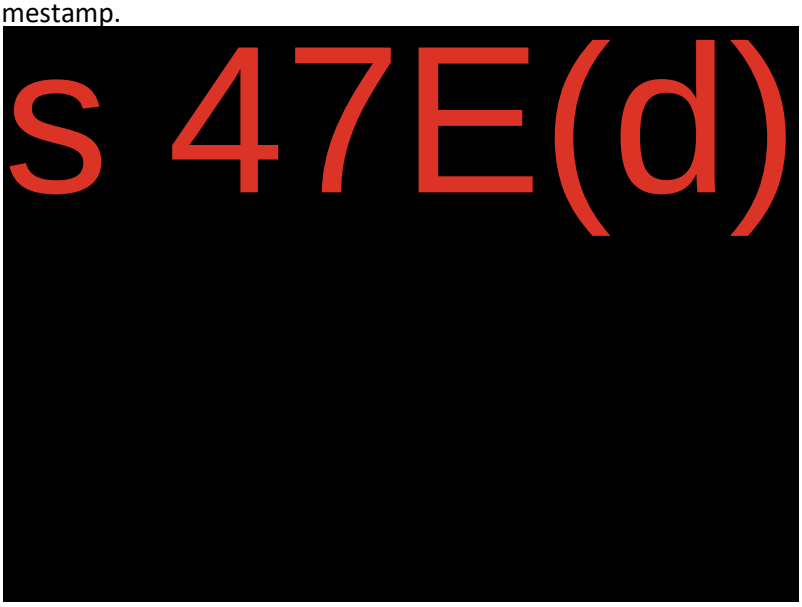
1. Use step 1 from above to reformat the dates to DD/MM/YYYY.
2. Highlight the date column and use the Find & Replace tool to remove any spaces.
3. If a timestamp appears alongside the date, use 'Text to columns' to remove.
    a. In the pop-up box that appears, make sure 'Fixed width' is selected.
    b. Select 'Next'.



4. In the 'Data preview' box, click to add a line between the end of the date and the timestamp.
    a. Select 'Next'.



5. In ‘Data preview’ select the column where the time is displayed.
    a. Select ‘Do not import column (skip)’ and then click ‘Finish’. This will remove the timestamp.



Reformat YYYY-MM-DD



1. Highlight the date column and click on the ‘Data’ tab. Select ‘Text to columns.’
    a. The below window will appear, select ‘Delimited’ and click ‘Next’.



2. Under the ‘Delimiters’ box, ensure there are no delimiters (tab, semicolon, comma etc) selected and click ‘Next’.





3. Select ‘Finish’. The dates should now be cleaned to the correct format of dd/mm/yyyy.

Addresses

Addresses appear either in a single cell or broken down into components such as street, suburb, state, etc. Cleaning an address focuses on ensuring the components of an address (street, state, suburb, etc) are in the correct columns. Note the steps below do not necessarily apply to international addresses, e.g., International postcodes will contain letters.

1. Review all address columns for non-address data. Methods for review include:
    a. Sort all columns by A-Z. This brings non-alphanumeric characters to the top.
    b. Run a length check in a new column using the LEN function. Filter column for outliers or anomalies based on the length (too short or long).
    c. Filter columns for ‘Contains’ = “@”.

   d. For Postcode columns: Sort Z-A or run a number check in a new column using a SEARCH function (Appendix A). Filter column for "FALSE" values to show cells containing non-numeric characters. Note a sort will only identify cells that begin with a non-numeric character, not cells that contain them, e.g., 1234abc.

2. Remove non-address data from column, relocate to correct column where possible.
3. Ensure the components of an address are in the correct column. Depending on how the data is delimited, addresses can be out of alignment, e.g., States in the Suburb column.
   a. For Suburbs, States, Countries and Postcodes: Use a length check to identify misplaced street names.
   b. For Streets, Suburbs, Countries and Postcodes: Use a SEARCH function or filters to check for state names.
   c. For Streets, Suburbs and Countries: Postcodes will appear at the top of the column when sorting A-Z.
   d. Use the filter drop-down to manually review the countries listed.
   e. Use the filter drop-down to manually review the suburbs listed. Check for 'Australia' in the suburb column as this occasionally appears here.
4. Relocate any cells in the incorrect column into the correct column.

Postcodes

Note that all Northern Territory postcodes begin with a 0. Excel will drop this 0 resulting in 3-digit postcodes. To add this back, use a BASE function =BASE(cell|10|4) in a manner similar to telephone cleaning processes described above.

Emails/Websites

Emails and websites are easy to identify as they contain an @ symbol or .com. It is important to check for emails and websites in incorrect columns as well as check for non-emails or multiple emails in the same cell.

1. Review email column for non-email data.
   a. Sort column by A-Z. This brings non-alphanumeric, followed by numeric characters to the top
   b. Filter column for 'Does Not Contain' = "@"
   c. Filter column for 'Contains'=".com"
2. Delete irrelevant data (Appendix B) and relocate other data into appropriate column.
3. Review for emails embedded amongst other data.
   a. Run a length check in a new column using the LEN function. Filter column for outliers or anomalies based on the length (too short or long).
4. Split emails from other data found in same cell.
   a. Manually split cell contents where there are minimal instances.
   b. Use a SUBSTITUTE and REPLACE function (Appendix A) to extract addresses.

5. Review for multiple email addresses (if not already revealed in the long end of the length check).
    a. Filter column for ‘Contains’ = “*@*@*” (two or more)

8. Ensure multiple email addresses (occurring in a single cell) are split into a new column or are separated by “;” depending on the circumstances. Columns should correspond to the correct entity, e.g., it will be assumed that ‘Email 2’ is linked to ‘Contact 2’. Concatenate with a “;” where there is more than one number for one entity.
    b. For minimal occurrences - Manually separate emails.
    c. For frequent or structured occurrences - Use the Find and Replace tool where a list has been separated by other characters. e.g., commas or spaces.

6. Move emails identified in incorrect columns into the email column. Concatenate with a “;” where there is more than one email. Note relocation becomes complicated by datasets with multiple email columns (primary or alternate).
    a. For minimal occurrences - Manually cut and paste two emails.
    b. For frequent or structured occurrences - Use a CONCAT function.

7. Remove hyperlinks
    a. Highlight column, right-click and select “Remove Hyperlinks”.

SEPERATING/CONCATENATING DATA

Separating Data

Text to columns is a tool that is used to separate data within a cell based on length or a specified character. Note - for unstructured cells this will not work.

=LEFT, =RIGHT, =MID functions will split cells based on a count of numbers from the left, right or specified middle position. This can be used in combination with a SEARCH function to split cells based on chosen delimiter – see example.

Separating Phone Numbers in Free Text Field

Below is an example of how to find and separate phone numbers in a data set. For further formulas to extract cell data see Appendix A.

1. Select cell B2, copy and paste the below formula into the formula bar.
    a. =MIN(IF(ISNUMBER(FIND({1,2,3,4,5,6,7,8,9,0}|A2))|FIND({1,2,3,4,5,6,7,8,9,0}|A2)))



    b. This will display the number of the position where the first number is located.
2. Use a LEFT and RIGHT function to display data to the left and right of the phone number.
    a. =LEFT(A2|B2-1) will display the data to the right of the number.



    b. =RIGHT(A2|LEN(A2)-B2+1) will display the number and all data to the left.



3. Use ‘Text to columns’ with ‘Fixed width’ separator. Note this can only be used on cells of the same structure.



4. Place two lines between the phone number and the text to the right. Hit ‘Next’.
    a. In the following pop-up box, select ‘Do not import (skip)’ when selecting the space between the number and email. This will ensure the space is removed.



5. Select ‘Finish’ and the cell will be separated into 2 columns.



Combining Data

Merging Cells

The **CONCAT** function allows you to combine data from multiple cells into one cell and allows you to add text in between.

=CONCAT(cell 1 to be combined|”;”|cell 2 to be combined)

When one cell has multiple entries (e.g., multiple phone numbers in one cell) and it is impractical or inefficient to create multiple columns, use “;” to separate the entries (e.g., “67891234;45678912”).

Merging Columns

One method of merging columns is to use the **XLOOKUP** function. This requires both sheets have at least one common identifier (e.g., both sheets have customer numbers). Note it may be better to upload sheets as separate datasets.

=XLOOKUP(value to search for|the array or range to search|array or range to return result from)

Note that there **must** be a unique search term (customer ID in the example below) for each row of data otherwise this formula will copy incorrect values.

Example – Using XLOOKUP to add outstanding payment information to the spreadsheet below.



1. In column H, enter the formula =XLOOKUP(B2|Sheet11!$A$1:$A$7|Sheet11!$B$1:$B$7).
    a. Note "$" ensure that the array stays the same when we duplicate this formula down the column. Clicking on the cells you want and dragging down ensures the correct arrays is selected.



2. Double-click or drag the black cross in the bottom right corner of the cell down to copy the formula down the column.



3. Copy and paste values only to replace the formula with the figures.

SAVING FILES

Formatting CSV Files

Files need to be ingested into [redacted] as a Comma-separated Value ("CSV") file.

Once the data has been cleaned, locally save a copy of the file as a .csv (comma delimited).

In Notepad, use the Replace tool to remove all spaces before or after a pipe (i.e. "| " to "|" and " |" to "|"). Note you cannot do " | " as this does not cover all instances (e.g., " |abc"). You can also use the tool to replace all double spaces with single spaces (" "to " ") if you forgot to do so in Excel.

Save the changes in Notepad.

[redacted]

Naming Conventions

Ensure that the correct naming conventions are utilised [redacted].

Raw Files

The raw file may need to be renamed to the convention if it is a combination of multiple files ([redacted]

Clean Files

[redacted]

CSV Files

[redacted]

Example

[redacted]

KEYBOARD SHORTCUTS

- Ctrl + Home (takes to top)
- Ctrl + End (takes to bottom/end of spreadsheet)
- Ctrl + Shift + End (highlight entire spreadsheet)
- Ctrl + *space* (highlight column)
- Shift + *space* (highlight row)
- Ctrl + shift + + (add row)
- Ctrl + - (delete row)
- Ctrl + 1 (format cells)
- Ctrl + j (in find and replace to make data linear)
- Alt + esv (paste special value)
- Alt + est (paste format)
- Ctrl + ` (displays the formula behind the data)
- Alt + = (displays the sum of numbers highlighted)

QUALITY ASSURANCE

QA Checklist

The Quality Assurance (QA) or 'sanity checking' phase is just as equally important as the data cleaning and ingestion component. The QA process ensures the data integrity is maintained and ensures that the hard work conducted so far allows for the information to be searchable within the organisation.

The QA process also includes new language to represent the roles of the Collection team throughout the QA process:

- **Analyst (Cleaner)-** This position is allocated to the Analyst who is involved in the cleaning and formatting phase of the data collection process (usually from the outset once data is received). The Analyst (Cleaner) will complete the 'Coversheet' tab and those sections marked:

  To be completed by the Analyst (Cleaner)

- **Analyst (Checker)-** This position is allocated to the Analyst who has the capacity to conduct the review process. The Analyst (Cleaner) must have an understanding of the data in its entirety to ensure that the relevant QA checks can be completed- the metadata and email chain can assist to contextualise the data. The checker completes the following categories: 'Modelling', 'Dataset', and 'Overall' sections and those sections marked:

  To be completed by the Analyst (Checker/Reviewer)

s 47E(d) [redacted]

[redacted]

s 47E(d) [redacted] o avoid confusion, the QA checklist is saved in a separate folder to the CLEAN and RAW versions. The QA checklist accompanies all data collections to be ingested into [redacted] (including Baseline and proactive collections). All sections of the document are to be filled out correctly and provided to a separate Analyst (Checker) within the collections team for quality assurance/sanity check.

A discussion between the Analyst (Cleaner) the data and the Analyst (Checker) **must** occur to ensure the sanity checklist is complied with as per the Data Cleaning Guidelines indicated above. Once discussed, an email **must** be sent by the Analyst (Cleaner) to the Analyst (Checker) and includes the clean and raw [redacted] links and any relevant cleaning notes made. This email **must** be sent from the 'Collections' Inbox s 47E(d) [redacted]

Any identified discrepancies must be rectified. Include the Assistant Director, Bulk Data Collections in the conversation if a solution cannot be reached or clarification sought.

***UPDATE 04 May 2026:** Collections Inbox will be changed to the Data Exploitation Inbox s 47E(d) [redacted] The Assistant Director, Bulk Data Collections will also be changed to Assistant Director, Data Insights Team 2.

Once the discrepancies are resolved (or if no issues were identified), the checker **must** return reply email to the 'Collections' inbox, with the Analyst (Cleaner) cc'd so that the email can be filed

appropriately. The Analyst (Cleaner) completes the final section of the document indicating that the data has been ingested into [redacted]

Tracker

The final stage of the process involves the inclusion onto the tracker [redacted]

The tracker is an accurate reflection of data the Bulk Data collection team has received and/or collected and collated, and responsible for the cleaning and ingestion of the data, formatted to be ingested into [redacted]

APPENDIX

Appendix A – Excel Functions

The following is a list of the more complex nested formulas that can be used in cleaning:

Number Check (FALSE value can be used to check for alphabet only values):
=IF(OR(ISNUMBER(SEARCH({"1","2","3","4","5","6","7","8","9","0"}|B2)))|"TRUE"|"FALSE")

Alpha Check (FALSE value can be used to check for numerical only values):
=IF(OR(ISNUMBER(SEARCH({"A","B","C","D","E","F","G","H","I","J","K","L","M","N","O","P","Q","R","S","T","U","V","W","X","Y","Z"}|B2)))|"TRUE"|"FALSE")

True/false check for Companies, States, Etc (check for common terms amongst data and add as needed): =SUMPRODUCT(--ISNUMBER(SEARCH({"*corp*","*trading*","*fund*","*bank*","*institute*","*organisation*","*foundations*","*industry*","*industries*","*organization*","*company*","*corporation*","* co *","* co. *","*ltd*","*group*","*university*","*trade*","*holding*","*directorate*","*bureau*","*force*","*centre*","*network*","*exchange*","*technology*","*committee*"}|C2)))>0

True/false check for Companies, States, Etc. Version 2

=SUMPRODUCT(--ISNUMBER(SEARCH({"*corp*","*trading*","*fund*","*bank*","*institute*","*organisation*","*foundations*","*industry*","*industries*","*organization*","*company*","*corporation*","*ltd*","*group*","*university*","*trade*","*holding*","*directorate*","*bureau*","*force*","*centre*","*network*","*exchange*","*technology*","*committee*","*association*","*council*","*academy*","*day care*","*district*","*early learning*","*services*","*college*","*society*","*community house*","*hospital*","*foundation*","*commission*","*school*","*consulting*","* co *","* co. *","* trust *","* pl *","* inc *","* inc. *","* shire *","* town *","* club *"}|F2)))>0

Note that "* xx *" format may not be picked up if xx is the last text in the cell with no trailing spaces. May need to do a "ends with" text filter for these ones.

Extract Email (Note will need to find & replace _*@*_ afterwards or use the REPLACE formula below to remove emails from original cell): =TRIM(IFERROR((RIGHT(SUBSTITUTE(LEFT(B2|FIND(" "|B2&" "|FIND("@"|B2))-1)|" "|REPT(" "|LEN(B2)))|LEN(B2)))|""))

Remove a string (text) (C2) after extracting it from its original cell (B2) (Note make sure said string has been converted into a text format: =TRIM(IFERROR(REPLACE(B2|FIND(C2|B2)|LEN(C2)|"")|""))

Extract Non-numeric (i.e. letters) to the left of the 1st number in a cell (beginning from left):
=LEFT(B2|MIN(FIND({0,1,2,3,4,5,6,7,8,9}|B2&"0123456789"))-1)

Extract Letters and numbers to the right of the 1st number in a cell (beginning from left):
=RIGHT(B2|LEN(B2)-MIN(FIND({0,1,2,3,4,5,6,7,8,9}|B2&"0123456789"))-1)

Extract Non-alpha characters (i.e. numbers) to the left of the 1st letter (beginning from left):
=LEFT(B2|MIN(SEARCH({"a","b","c","d","e","f","g","h","i","j","k","l","m","n","o","p","q","r","s","t","u","v","w","x","y","z"}|B2&"abcdefghijklmnopqrstuvwxyz"))-1)

Extract numbers and letters to the right of the 1st letter in a cell (beginning from left):
=RIGHT(B2|LEN(B2)-
MIN(SEARCH({"a","b","c","d","e","f","g","h","i","j","k","l","m","n","o","p","q","r","s","t","u","v","w","x","y","z","("}|B2&"abcdefghijklmnopqrstuvwxyz("))+1)

Formula to replace "Alt Enter" (i.e., line breaks) with "Space"

=TRIM(SUBSTITUTE(SUBSTITUTE(**A2**|CHAR(13)|"")|CHAR(10)|" "))

Replace the content in the last " " to replace alt enter with something else

Appendix B – Irrelevant Data

Irrelevant data refers to any cell with entirely immaterial data. These cells are found in datasets across the board. Examples include but cannot be limited to:

- NULL
- NONE
- N/A
- *
- -
- 0
- &

Appendix C – Phone Number List Separators

Multiple phone numbers in a single cell are often split by the following characters (separators). These can be used to identify the end of one phone number. Note spaces cannot be relied upon as there are often spaces between numbers themselves, e.g., 0412 345 678.

- Work
- Home
- Hm
- Wk
- H
- W
- Fax
- F
- ()
- &
- ,

Appendix D – Keywords used to identify Company Names

The following is a list of words commonly associated with corporations and organisations:

- Corp
- “ Co “
- “ co.”
- Org
- Pty
- Ltd
- Group
- Institute
- Fund
- Bank
- Foundations
- Industry
- Industries
- Holding
- Directorate
- Bureau
- Force
- Centre
- Network
- Exchange
- Technology
- Committee
- Trust
- “ inc”
- Limited
- Partners
- Proprietary
- investment

Appendix E – Keywords used to identify Multiple Names

- “ and “
- /
- &
- ,

VERSION CONTROL

IMPLEMENTATION DATE	VERSION	AUTHOR	s 47E(d)	NOTES
s 47E(d)				

IMPLEMENTATION DATE	VERSION	AUTHOR	s 47E(d)	NOTES
s 47E(d)	[redacted]	[redacted]	[redacted]	[redacted]

Alterations to this document should not be made without prior approval and consultation with the Assistant Director- Data Insights Team 2.

Document 6

AUSTRALIAN CRIMINAL INTELLIGENCE COMMISSION

Model Cross Check

To be completed by the Analyst (Cleaner)

Column Heading	Domain	[1]	[2]	[3]	[4]	[5]	[6]
s 37(2)(b), s 47E(d)		x	x	x	x	x	
		x	x	x	x		
		x					
			x				
				x			
					x		
		x	x	x	x	x	
		x	x	x	x	x	
						x	
						x	
						x	
						x	
						x	
						x	
						x	
						x	

s 37(2)(b), s 47E(d)

Suggested alterations to be marked with a red X.

Quality Assurance/Sanity Checking

To be completed by the Analyst (Checker/Reviewer)

Modelling	YES, NO, N/A	Comments
Has a review of the above Model been completed?	s 47E(d)	
Is the Modelling correct and accurate in line with the Cleaning Guidelines? *If no, speak to Analyst (Cleaner) first and/or consult Data Cleaning guidelines* s 47E(d)		

Dataset	YES, NO, N/A	Comments
Are filters and hidden columns turned off?	s 47E(d)	o be turned off
Has a check been completed to remove any 'double spaces' from the rows and columns?		
Has a check of the filter options (for abnormalities) been conducted? e.g. postcodes with 3 digits instead of 4 etc. *If there are abnormalities identified, speak to Analyst (Cleaner) first and/or consult Data Cleaning guidelines-* s 47E(d)		
Is the data within the columns formatted correctly in line with the Cleaning Guidelines?		
Are column headings worded correctly in line with the Cleaning Guidelines?		
Are columns seperated appropriately in line with the Cleaning Guidelines? e.g., phone numbers, emails etc		
Has a check of the top, middle and and bottom data rows been completed against the raw data in line with the Cleaning Guidelines? *(to be completed after data has been ordered)*		
Have 5 'random' rows (minimum) been checked against the raw data and are correct in line with the Cleaning Guidelines?		
Does the data represent its truest form as received? *Review both the Clean and Raw data in its entirety*		

Overall	YES, NO, N/A	Comments
Has this request come to the Analyst (Checker/Reviewer) via email with the s 47E(d) for the Clean/Raw version?	s 47E(d)	
Has a return email been sent to the Collections Inbox indicating the Quality Assurance has been completed? *(email to be filed in appropriate email folder)*		
Is the Analyst (Checker/Reviewer) satisfied that the data is cleaned correctly with the correct Modelling?		
Is the data ready for ingestion?		
Has a check of the file name (the s 47E(d) to the raw file) been completed and is named correctly in line with the Cleaning Guidelines?		

Upload to s 47E(d) when the above checklist is completed and all discrepancies rectified.

To be completed by the Analyst (Cleaner)

After Upload- Next Day		Comments
Has a check of the Modelling been conducted in s 47E(d) to ensure nothing has dropped off? *(check the indexing in s 47E(d) to ensure it is consistent with the above Modelling- the dots)*	s 47E(d)	
Has the data been ingested into s 47E *(Using s 47E(d), s 37(2)(b) , conduct a search utilising a unique name or feature if possible. If the related data is located, it is ingested correctly and is searchable)*		
Has the data been semantically mapped? *(Select no if the data is unable to be semantically mapped, otherwise, this needs to be completed)*		

Document 7

Tab 1.1

Collection Requester	Determination	Collection Target Type	Process	Notice No. (if reqd.)	Target	Collection Coordinator (responsible for s.21a Notice preparation etc.)	Collector (responsible for service etc.)
FFT	SFC 2022	Financial Crime	s 37(2)(b), s 47E(d)			s 22(1)(a)(ii)	
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					
FFT	SFC 2022	Financial Crime					

Tab 1.2

Date of Notice Issue or Request Date	Date Served	Date of collection	FDH Ingestion Date (if applicable)	Ingestion Completed by	No. of Rows Received	Data s 47E(d) Reference	Comments (meeting notes, collection status, contacts etc.)	Semantic Mapping Complete
s 47E(d)				s 22(1)(a)(ii)			s 47E(d)s 37(2)(b), s 47E(d)	

Tab 2.1

Collection Requester	Determination	Collection Target Type	Process	Notice No. (if reqd.)
FFT	SFC 2022	Financial Crime	s 37(2)(b), s 47E(d)	
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT		Imports		
FFT		Imports		
FFT	Drugs 2022 SO	Other		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		
FFT	SFC 2022	Financial Crime		

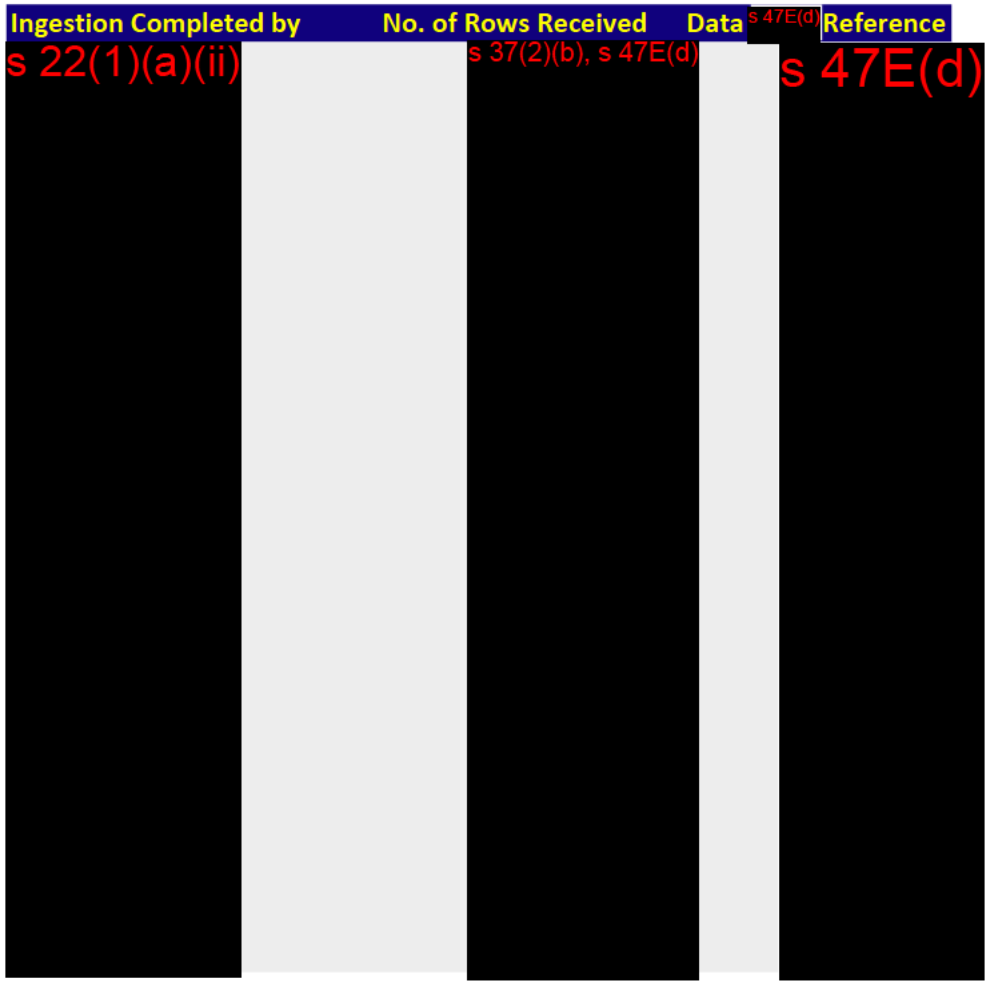
Tab 2.2

Target	Collection Coordinator (responsible for s.21a Notice preparation etc.)
s 37(2)(b), s 47E(d)	s 22(1)(a)(ii)

Tab 2.3

Collector (responsible for service etc.)	Date of Notice Issue or Request Date	Date Served	Date of collection	FDH Ingestion Date (if applicable)
s 22(1)(a)(ii)				

Tab 2.4


Ingestion Completed by
No. of Rows Received
Data
s 47E(d)
Reference
s 22(1)(a)(ii)
s 37(2)(b), s 47E(d)
s 47E(d)

Tab 2.5

Comments (meeting notes, collection status, contacts etc.)	Semantic Mapping Complete
s 37(2)(b), s 47E(d)	s 37(2)(b), s 47E(d)

Tab 3.1

Who is the Data Custodian? (Position Title- Not a person- Position must be EL2) ACIC INTERNAL	Who is the Point of Contact? (Group Inbox Only e.g. Collections@acic.gov.au) ACIC INTERNAL	Collection Coordinator (Person responsible and Notice s.21a Notice preparation etc.)	ACIC Determination (Drop Down)	Project Name	ACIC National Criminal Intelligence Priorities (NCIP) (Drop Down)	Collection Type	Notice No. (if applicable)	Which Agency Disseminated this to the ACIC? EXTERNAL/INTERNAL	Collector (responsible for service etc.)	Date Notice Issued or Request Date	Date Notice was Served
Director, Serious Fraud	s 47E(d)	s 22(1)(a)(ii)	SFC 2022	s 47E(d), s 37(2)(b	Serious Financial Crimes	s 37(2)(b), s 47E(d)			s 22(1)(a)(ii)		s 37(2)(b), s 47E(d)
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Fraud Control & Analytics			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2022		Serious Financial Crimes						
Director, Serious Fraud			SFC 2025		Serious Financial Crimes						
Director, Serious Fraud			SFC 2025		Serious Financial Crimes						
Director, Serious Fraud			SFC 2025		Serious Financial Crimes						

Tab 3.2

Date received by ACIC (Date of Collection)	Date received by Collections team	Ingestion Location (Drop Down)	Date Ingested into the [redacted] (if applicable)	Ingestion Completed by (Drop Down)	No. of Rows Received	Data [redacted] Reference (Hyperlink [redacted] link)	What is the data? (Succinct Description)	Semantic Mapping Complete? (Drop Down)
s 37(2)(b), s 47E(d)				s 22(1)(a)(ii)	s 37(2)(b), s 47E	s 47E(d)	s 37(2)(b), s 47E(d)	

Tab 4.1

Who is the Data Custodian? (Position Title- Not a person- Position must be EL2) ACIC INTERNAL	Who is the Point of Contact? (Group Inbox Only e.g. Collections@acic.gov.au) ACIC INTERNAL	Collection Coordinator (Person responsible and Notice s.21a Notice preparation etc.)	ACIC Determination (Drop Down)	Project Name	ACIC National Criminal Intelligence Priorities (NCIP) (Drop Down)	Collection Type	Notice No. (if applicable)	Which Agency Disseminated this to the ACIC? EXTERNAL/INTERNAL	Collector (responsible for service etc.)
Director, Serious Fraud	s 47E(d)	s 22(1)(a)(ii)	SFC 2025		Serious Financial Crimes	s 37(2)(b), s 47E(d)			s 22(1)(a)(ii)
Director - Serious Fraud			SFC 2025		Serious Financial Crimes				
Director - Serious Fraud			SFC 2025		Serious Financial Crimes				
Director - Serious Fraud			SFC 2025		Serious Financial Crimes				
Director, Serious Fraud			SFC 2025	s 37(2)(b),	Serious Financial Crimes				

Tab 4.2

Date Notice Issued or Request Date	Date Notice was Served	Date received by ACIC (Date of Collection)	Date received by Collections team	Ingestion Location (Drop Down)	Date Ingested into the FDH (if applicable)	Ingestion Completed by (Drop Down)	No. of Rows Received	Data Reference (Hyperlink link)	What is the data? (Succinct Description)	Semantic Mapping Complete? (Drop Down)
s 37(2)(b), s 47E(d)					s 47E(d)	s 22(1)(a)(ii)	s 37(2)(b), s 47E(d)	s 47E(d)	s 37(2)(b), s 47E(d)	

Document 8

From: s 22(1)(a)(ii)
To: s 22(1)(a)(ii); s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: s 47E(d), s 37(2)(b); s 47E(d), s 37(2)(b) [SEC=OFFICIAL:Sensitive]
Date: Tuesday, 7 July 2026 3:53:00 PM
Attachments: image001.png
image002.png
image003.png
image004.png
image005.png
image006.jpg
image007.png
image008.png
image009.png
image010.jpg

OFFICIAL: Sensitive

Thanks s 22(1)(a)

Appreciate you reaching out.

Just confirm the Point of Contact Email- that more so relates to communication internally, so if you have a distribution list or team Inbox, this is what we're after.

Much appreciated.

s 22(1)(a)

s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Tuesday, 7 July 2026 3:25 PM
To: s 22(1)(a)(ii)
Cc: s 22(1)(a)(ii)
Subject: s 47E(d), s 37(2)(b); s 47E(d)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)

I'll take this one – please see below updates.

Determination and Project the data was collected under:
Organisation Collected from: s 47E(d)
Description of the Data: s 37(2)(b), s 47E(d)
Request date: s 47E(d), s 37(2)(b)
Date received by the ACIC: 03 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 47E(d)
Date range of the data: Current as of 31 March 2026
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 47E(d)

s 47E(d)

Data custodian (EL2 position from unit): Director - Serious Fraud
Point of Contact (communal email- not person): s 47E(d)



s 22(1)(a)(ii)
Intelligence Analyst | Serious Fraud
Operational Analysis Branch
Australian Criminal Intelligence Commission
s 22(1)(a)(ii)

First Nation artwork

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future.

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 11:14 AM
To: s 22(1)(a)(ii)
Cc: Data Exploitation s 47E(d) s 22(1)(a)(ii)
Subject: s 47E(d), s 37(2)(b) [SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)(ii)

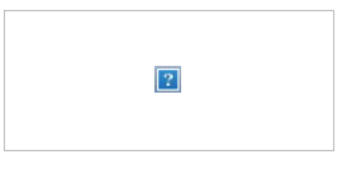
Thanks for sending this through.

This looks good to ingest into the s 47E(d) so will get the analysts to review this for ingestion.

FFC team, can you please confirm and fill in the remainder of the meta data tags for us please?

Determination and Project the data was collected under:
Organisation Collected from: s 37(2)(b), s 47E(d)
Description of the Data:
Request date: Ongoing
Date received by the ACIC: 06 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 37(2)(b), s 47E(d)
Date range of the data:
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 47E(d)
Data custodian (EL2 position from unit): Director- Serious Fraud
Point of Contact (communal email- not person):

s 22(1)(a)



s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 8:10 AM
To: Data Exploitation s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: FW: s 47E(d), s 37(2)(b)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

FYI

From: s 47E(d) Fraud Fusion Taskforce s 47E(d)
Sent: Friday, 3 July 2026 2:28 PM
To: FraudFusionCentre s 47E(d) ; s 22(1)(a)(ii)
Cc: s 47E(d) - Fraud Fusion Taskforce s 47E(d)
Subject: s 37(2)(b), s 47E(d) [SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

G'day ACIC FFC team,

The following information is disclosed under s 47E(d)

To ensure the s 47E(d) maintains updated data holdings in the ACIC FFC, I am proactively disclosing the s 37(2)

Please find attached:

- An Excel dataset – s 37(2)(b), s 47E(d)
-
- which gives further information on permitted use and disclosure.

Please contact s 47E(d) if you have any questions.

s 47E(d), s 47F(1)

OFFICIAL: Sensitive

Notice:
The information contained in this email message and any attached files may be confidential information, and may also be the subject of legal professional privilege. If you are not the intended recipient, any use, disclosure or copying of this email is unauthorised. If you received this email in error, please notify the sender by contacting the department's switchboard on 1300 566 046 during business hours (8:30am - 5pm Canberra time) and delete all copies of this transmission together with

any attachments.

Document 9

From: s 22(1)(a)(ii)
To: Data Exploitation; s 22(1)(a)(ii)
Cc: s 22(1)(a)(ii)
Subject: FOR TASKING: s 37(2)(b), s 47E(d) [SEC=OFFICIAL:Sensitive]
Date: Thursday, 9 July 2026 8:48:00 AM
Attachments: image001.png
image002.png
image003.png
image004.png
image005.png
image006.jpg
image007.png
image008.png
image009.png
image010.jpg

OFFICIAL: Sensitive

Hi s 22(1)(d)

I was able to find the best email address, so I've updated it below:

Determination and Project the data was collected under: s 47E(d), s 37(2)(b)
Organisation Collected from: s 47E(d)
Description of the Data: s 37(2)(b), s 47E(d)
Request date: s 47E(d)
Date received by the ACIC: 03 July 2026
Method of collection: s 37(2)(b), s
CM link to the original data: s 37(2)(b), s 47E(d)
Date range of the data: Current as of 31 March 2026
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 37(2)(b), s 47E(d)
Data custodian (EL2 position from unit): Director - Serious Fraud
Point of Contact (communal email- not person): s 47E(d)

@Data Exploitation- for completion.

s 22(1)(a)

s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

From: s 22(1)(a)(ii)
Sent: Tuesday, 7 July 2026 3:53 PM
To: s 22(1)(a)(ii) Data Exploitation s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: RE: s 37(2)(b), s 47E(d)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Thanks s 22(1)(a)

Appreciate you reaching out.

Just confirm the Point of Contact Email- that more so relates to communication internally, so if you have a distribution list or team Inbox, this is what we're after.

Much appreciated.

s 22(1)(a)



s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Tuesday, 7 July 2026 3:25 PM
To: s 22(1)(a)(ii)
Cc: s 22(1)(a)(ii)
Subject: RE: s 37(2)(b), s 47E(d)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)

I'll take this one – please see below updates.

Determination and Project the data was collected under:
Organisation Collected from: s 47E(d)
Description of the Data: s 37(2)(b), s 47E(d)

Request date: s 37(2)(b), s 47E(d)
Date received by the ACIC: 03 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 37(2)(b), s 47E(d)
Date range of the data: Current as of 31 March 2026
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 37(2)(b), s 47E(d)
Data custodian (EL2 position from unit): Director - Serious Fraud
Point of Contact (communal email- not person): s 47E(d)



s 22(1)(a)(ii)
Intelligence Analyst | Serious Fraud
Operational Analysis Branch
Australian Criminal Intelligence Commission
s 22(1)(a)(ii)

First Nation artwork



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future.

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 11:14 AM
To: s 22(1)(a)(ii)
Cc: Data Exploitation s 47E(d) s 22(1)(a)(ii)
Subject: s 47E(d), s 37(2)(b) [SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)(ii)

Thanks for sending this through.

This looks good to ingest into the s 47E(d), so will get the analysts to review this for ingestion.

FFC team, can you please confirm and fill in the remainder of the meta data tags for us please?

Determination and Project the data was collected under:
Organisation Collected from: s 47E(d)
Description of the Data:
Request date: Ongoing
Date received by the ACIC: 06 July 2026
Method of collection: s 37(2)(b), s 47E(d)
CM link to the original data: s 37(2)(b), s 47E(d)
Date range of the data:
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 47E(d), s 37(2)(b)
Data custodian (EL2 position from unit): Director- Serious Fraud
Point of Contact (communal email- not person):

s 22(1)(a)(ii)



s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 8:10 AM
To: Data Exploitation s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: FW: s 47E(d), s 37(2)(b)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

FYI

From: s 47E(d) Fraud Fusion Taskforce s 47E(d)
Sent: Friday, 3 July 2026 2:28 PM
To: FraudFusionCentre s 47E(d) ; s 22(1)(a)(ii)
Cc: s 47E(d) Fraud Fusion Taskforce s 47E(d)
Subject: s 37(2)(b), s 47E(d) [SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

G'day ACIC FFC team,

The following information is disclosed under s 47E(d)

To ensure s 47E(d) maintains updated data holdings in the ACIC FFC, I am proactively disclosing the s 47E(d)

Please find attached:

- An Excel dataset – s 37(2)(b), s 47E(d)
-
- which gives further information on permitted use and disclosure.

Please contact s 47E(d) if you have any questions.

s 37E(d), s 47F(1)

OFFICIAL: Sensitive

Notice:
The information contained in this email message and any attached files may be confidential information, and may also be the subject of legal professional privilege. If you are not the intended recipient, any use, disclosure or copying of this email is unauthorised. If you received this email in error, please notify the sender by contacting the department's switchboard on 1300 566 046 during business hours (8:30am - 5pm Canberra time) and delete all copies of this transmission together with any attachments.

Document 10

From: s 22(1)(a)(ii)
To: Data Exploitation; s 22(1)(a)(ii)
Cc: s 22(1)(a)(ii)
Subject: FOR TASKING: s 47E(d), s 37(2)(b) [SEC=OFFICIAL:Sensitive]
Date: Thursday, 9 July 2026 8:48:00 AM
Attachments: image001.png
image002.png
image003.png
image004.png
image005.png
image006.jpg
image007.png
image008.png
image009.png
image010.jpg

OFFICIAL: Sensitive

Hi s 22(1)(a)(ii)

I was able to find the best email address, so I've updated it below:

Determination and Project the data was collected under: s 37(2)(b), s 47E(d)
Organisation Collected from: s 47E(d)
Description of the Data: s 37(2)(b), s 47E(d)
Request date: s 37(2)(b), s 47E(d)
Date received by the ACIC: 03 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 37(2)(b), s 47E(d)
Date range of the data: Current as of 31 March 2026
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 47E(d), s 37(2)(b)
Data custodian (EL2 position from unit): Director - Serious Fraud
Point of Contact (communal email- not person): s 47E(d)

@Data Exploitation- for completion.

s 22(1)(a)(ii)

s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

From: s 22(1)(a)(ii)
Sent: Tuesday, 7 July 2026 3:53 PM
To: s 22(1)(a)(ii) Data Exploitation s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: RE: s 47E(d), s 37(2)(b)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Thanks s 22(1)(a)(ii)

Appreciate you reaching out.

Just confirm the Point of Contact Email- that more so relates to communication internally, so if you have a distribution list or team Inbox, this is what we're after.

Much appreciated.

s 22(1)(a)



s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Tuesday, 7 July 2026 3:25 PM
To: s 22(1)(a)(ii)
Cc: s 22(1)(a)(ii)
Subject: RE: s 47E(d), s 37(2)(b)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)

I'll take this one – please see below updates.

Determination and Project the data was collected under:
Organisation Collected from: s 47E(d)
Description of the Data: s 37(2)(b), s 47E(d)

Request date: s 37(2)(b), s 47E(d)
Date received by the ACIC: 03 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 37(2)(b), s 47E(d)
Date range of the data: Current as of 31 March 2026
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 37(2)(b), s 47E(d)
Data custodian (EL2 position from unit): Director - Serious Fraud
Point of Contact (communal email- not person): s 47E(d)



s 22(1)(a)(ii)
Intelligence Analyst | Serious Fraud
Operational Analysis Branch
Australian Criminal Intelligence Commission
s 22(1)(a)(ii)

First Nation artwork



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future.

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 11:14 AM
To: s 22(1)(a)(ii)
Cc: Data Exploitation s 47E(d) s 22(1)(a)(ii)
Subject: s 47E(d), s 37(2)(b) [SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

Hi s 22(1)(a)(ii)

Thanks for sending this through.

This looks good to ingest into the s 47E(d) so will get the analysts to review this for ingestion.

FFC team, can you please confirm and fill in the remainder of the meta data tags for us please?

Determination and Project the data was collected under:
Organisation Collected from: s 47E(d)
Description of the Data:
Request date: Ongoing
Date received by the ACIC: 06 July 2026
Method of collection: s 37(2)(b), s 47E(d)
link to the original data: s 37(2)(b), s 47E(d)
Date range of the data:
Security Classification: OFFICIAL: Sensitive
Caveats on the data: s 37(2)(b), s 47E(d)
Data custodian (EL2 position from unit): Director- Serious Fraud
Point of Contact (communal email- not person):

s 22(1)(a)(ii)



s 22(1)(a)(ii)
Assistant Director | Data Insights Team 2
Data Analytics Section | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Gadigal Country
s 22(1)(a)(ii)



I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

---

From: s 22(1)(a)(ii)
Sent: Monday, 6 July 2026 8:10 AM
To: Data Exploitation s 47E(d)
Cc: s 22(1)(a)(ii)
Subject: FW: s 47E(d), s 37(2)(b)
[SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

FYI

From: [redacted] - Fraud Fusion Taskforce [redacted]
Sent: Friday, 3 July 2026 2:28 PM
To: FraudFusionCentre [redacted]
Cc: [redacted] Fraud Fusion Taskforce [redacted]
Subject: [redacted] SEC=OFFICIAL:Sensitive]

OFFICIAL: Sensitive

G'day ACIC FFC team,

The following information is disclosed under [redacted]

To ensure the [redacted] maintains updated data holdings in the ACIC FFC, I am proactively disclosing the [redacted]
Please find attached:

- An Excel dataset – [redacted]
- [redacted] We can provide details if required.
- [redacted] which gives further information on permitted use and disclosure.

Please contact [redacted] if you have any questions.

[redacted]

OFFICIAL: Sensitive

Notice:
The information contained in this email message and any attached files may be confidential information, and may also be the subject of legal professional privilege. If you are not the intended recipient, any use, disclosure or copying of this email is unauthorised. If you received this email in error, please notify the sender by contacting the department's switchboard on 1300 566 046 during business hours (8:30am - 5pm Canberra time) and delete all copies of this transmission together with any attachments.

Document 11

From: s 22(1)(a)(ii)
To: Data Exploitation
Cc: s 22(1)(a)(ii)
Subject: INGESTION: s 37(2)(b), s 47E(d) [SEC=OFFICIAL:Sensitive]
Date: Wednesday, 15 July 2026 9:13:31 AM
Attachments: image001.png
image002.png
image003.png
image004.png
image005.png
image006.png
image007.jpg

OFFICIAL: Sensitive

Hi everyone,

The s 37(2)(b), s 47E(d) data have been uploaded to s 47E(d) and is now searchable in Palantir.

Please see the links below for the uploaded CSV files, as well as the Clean and Raw files.

s 37(2)(b), s 47E(d)	**CSV** s 37(2)(b), s 47E(d)
s 37(2)(b), s 47E(d)	**CSV** s 37(2)(b), s 47E(d)
s 37(2)(b), s 47E(d)	**CSV** s 37(2)(b), s 47E(d)

CLEAN File: s 37(2)(b), s 47E(d)

RAW file: s 37(2)(b), s 47E(d)

s 47E(d), s 37(2)(b)
s 37(2)(b), s 47E(d)

Thanks and regards,

s 22(1)(a)(ii)

s 22(1)(a)(ii)
Intelligence Analyst
Data Insights Team 1 | Strategic Analysis & Outreach Branch
Australian Criminal Intelligence Commission | Ngunnawal Country
s 22(1)(a)(ii)

First Nation artwork

I acknowledge the Traditional Custodians of the lands and waters where we live and work, and pay my respects to Elders past, present and future

Document 12

	Column	Name	Domain	[1]	[2]	[3]	[4]
	1	s 47E(d), s 37(2)(b)		●	○	○	○
	2			●	○	○	○
	3			●	○	○	○
	4			●	●	●	●
	5			●	○	○	○
	6			●	○	○	○
	7			●	○	○	○
	8			●	○	○	○
	9			●	○	○	○
	10			○	●	●	●
	11			○	●	○	○
	12			○	○	●	○
	13			○	○	○	●
	14			●	●	●	●
	15			●	●	●	●
	16			●	○	○	○
	17			●	●	●	●
	18			●	○	○	○
	19			●	○	○	○
	20			●	○	○	○

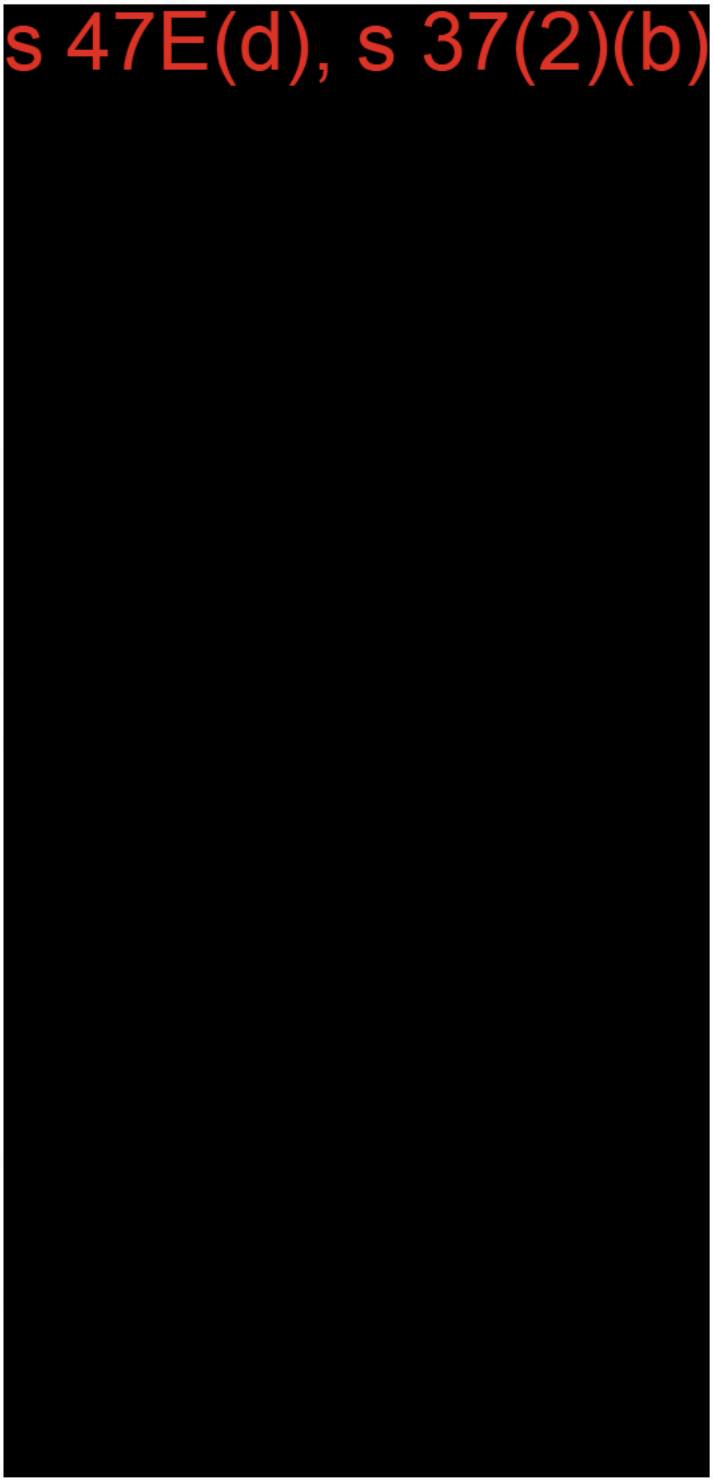
Document 13

	Column	Name	Domain	[1]	[2]	[3]	[4]
	1	s 47E(d), s 37(2)(b)	[redacted]	■	■	■	■
	2			■	■	■	□
	3			■	□	□	□
	4			□	■	□	□
	5			□	□	■	□
	6			■	■	■	■
	7			■	■	■	■
	8			□	□	□	■
	9			■	■	■	■
	10			□	□	□	■
	11			□	□	□	■
	12			□	□	□	■
	13			□	□	□	■
	14			□	□	□	□
	15			□	□	□	□
	16			□	□	□	□
	17			□	□	□	■
	18			□	□	□	□
	19			□	□	□	□
	20			□	□	□	□
	21			□	□	□	□
	22			□	□	□	□
	23			□	□	□	□
	24			□	□	□	□
	25			□	□	□	□
	26			□	□	□	□
	27			□	□	□	□
	28			□	□	□	□
	29			□	□	□	□
	30			□	□	□	□
	31			□	□	□	□
	32			□	□	□	□
	33			□	□	□	□
	34			□	□	□	□
	35			□	□	□	□
	36			□	□	□	□
	37			□	□	□	□
	38			□	□	□	□
	39			□	□	□	□
	40			□	□	□	□
	41			□	□	□	□
	42			□	□	□	□
	43			□	□	□	□
	44			□	□	□	□

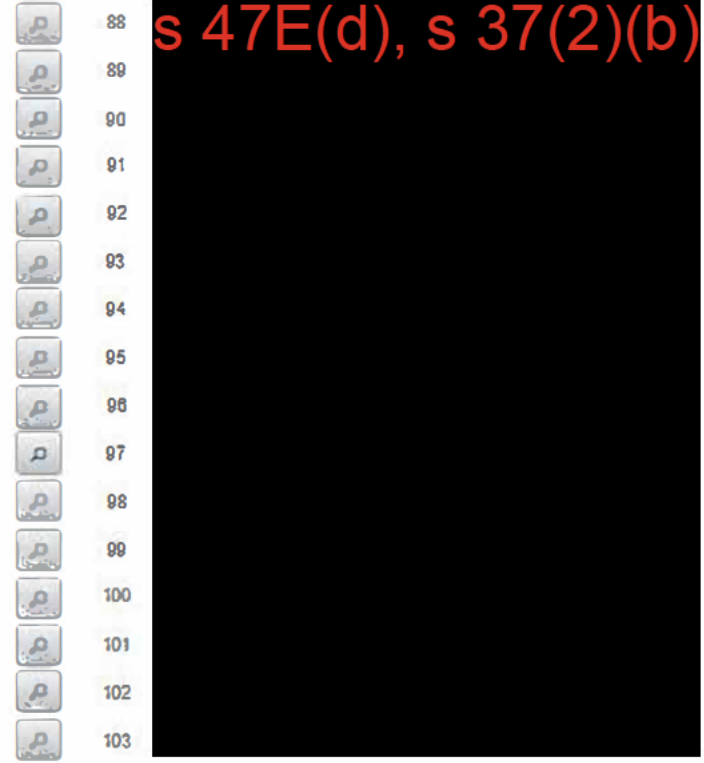
s 47E(d), s 37(2)(b)

Document 14

Column	Name	Domain	[1]	[2]	[3]	[4]	[5]
1	s 47E(d), s 37(2)(b)		●	●	●	●	●
2			●	●	●	●	○
3			●	○	○	○	○
4			○	●	○	○	○
5			○	○	●	○	○
6			○	○	○	●	○
7			●	●	●	●	●
8			●	●	●	●	●
9			○	○	○	○	●
10			○	○	○	○	●
11			○	○	○	○	●
12			○	○	○	○	●
13			○	○	○	○	●
14			○	○	○	○	●
15			○	○	○	○	●
16			○	○	○	○	●
17			○	○	○	○	●
18			○	○	○	○	○
19			○	○	○	○	○
20			○	○	○	○	○
21			○	○	○	○	○
22			○	○	○	○	○
23			○	○	○	○	○
24			○	○	○	○	○
25			○	○	○	○	○
26			○	○	○	○	○
27			○	○	○	○	○
28			○	○	○	○	○
29			○	○	○	○	○
30			○	○	○	○	○


s 47E(d), s 37(2)(b)


s 47E(d), s 37(2)(b)


s 47E(d), s 37(2)(b)

Getting Started with Palantir Gotham

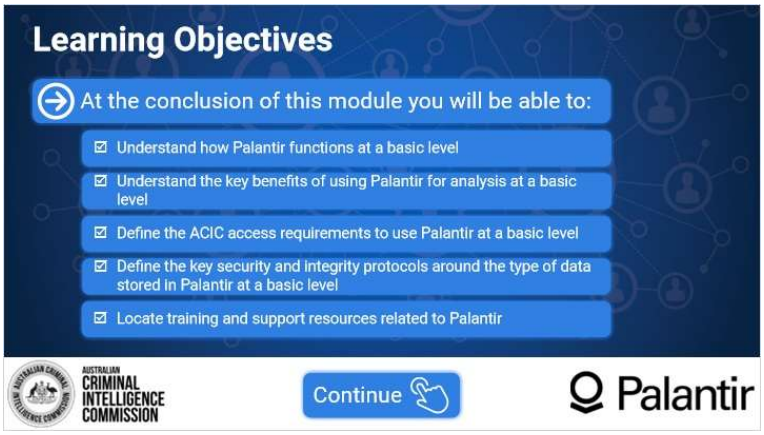
1. Module 1: Getting Started

1.1 Palantir L-Plates



Notes:

1.2 Learning Objectives



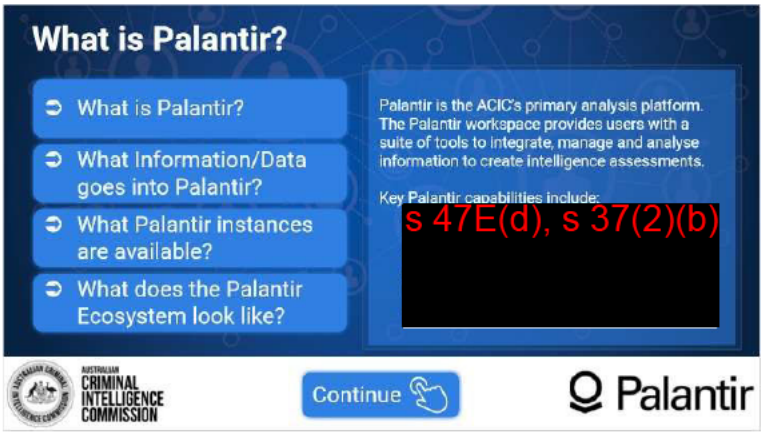
1.3 Navigating this Module



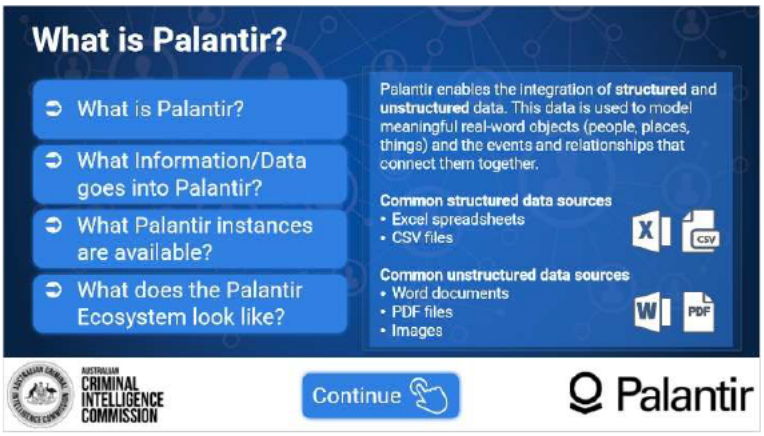
1.4 What is Palantir?



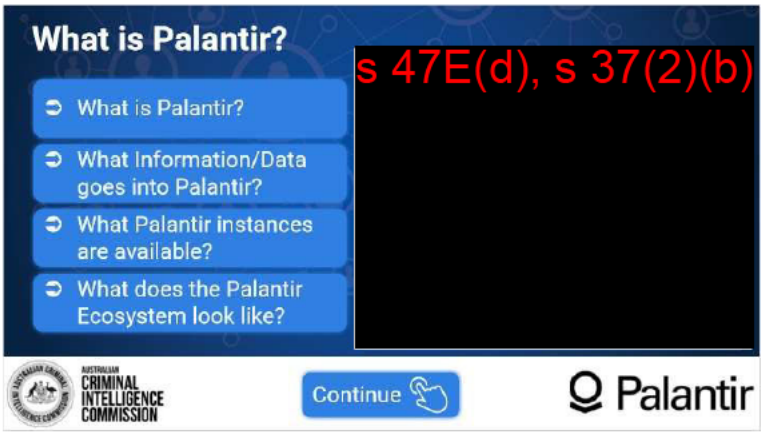
What is Palantir (Slide Layer)



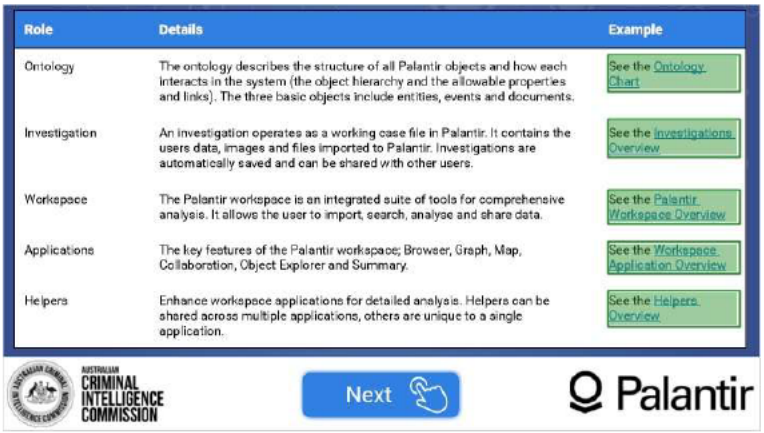
Information/data (Slide Layer)



Palantir Architecture (Slide Layer)

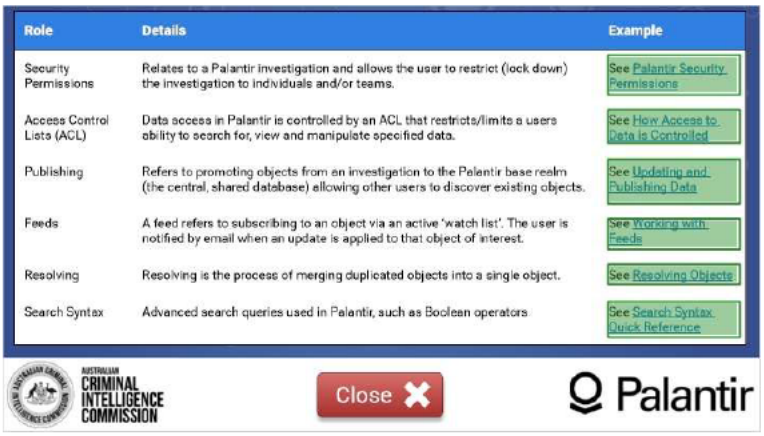


Palantir Ecosystem 1 (Slide Layer)



Published by Articulate® Storyline www.articulate.com

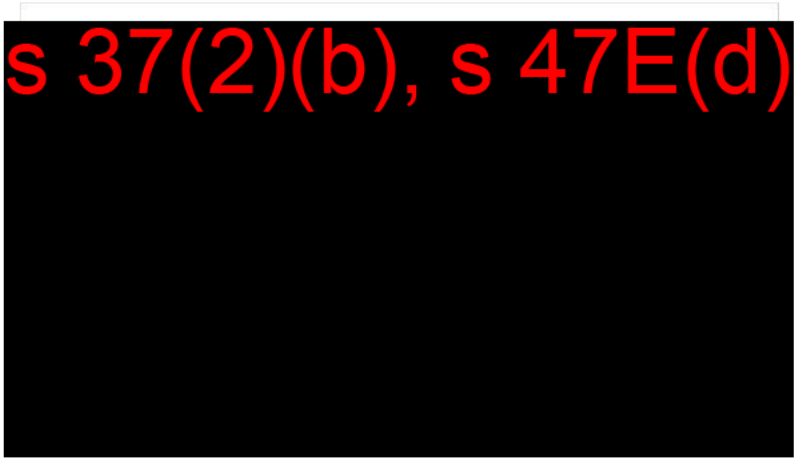
Palantir Ecosystem 2 (Slide Layer)



1.5 Why do we use Palantir?

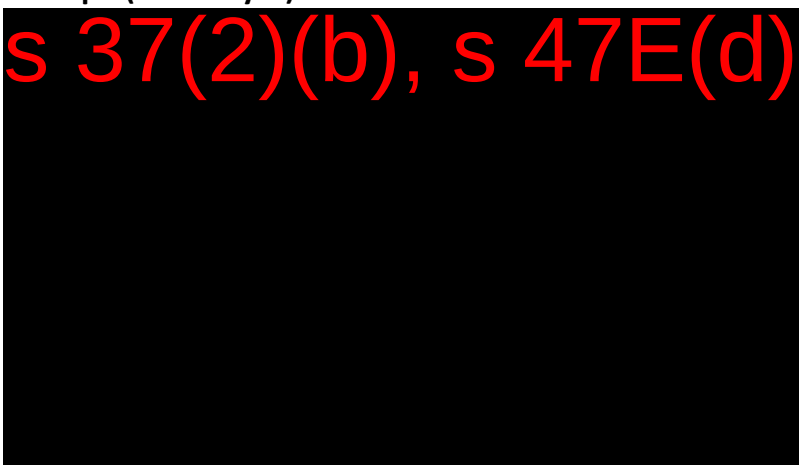


Browser (Slide Layer)

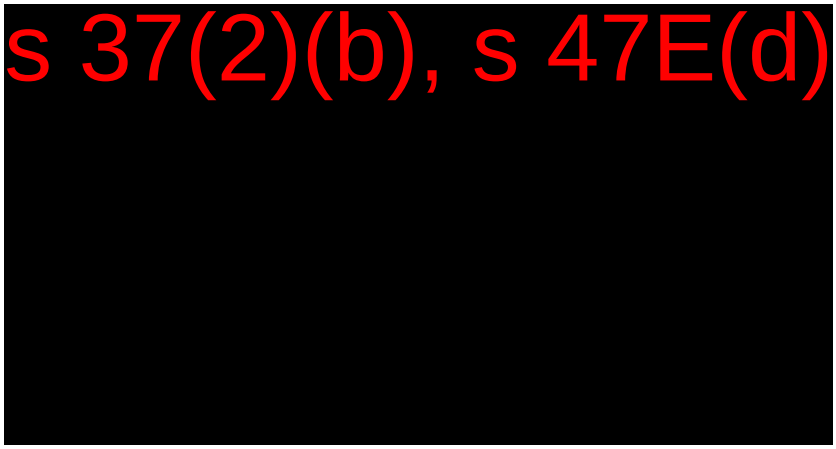


Published by Articulate® Storyline www.articulate.com

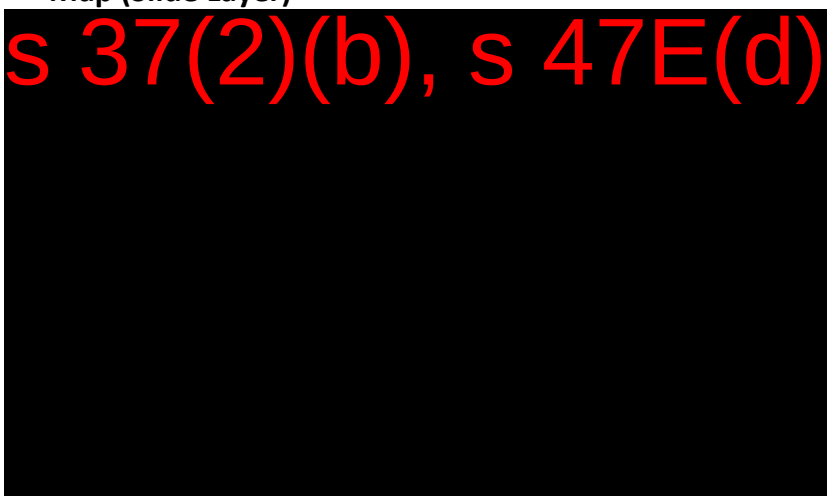
Graph (Slide Layer)



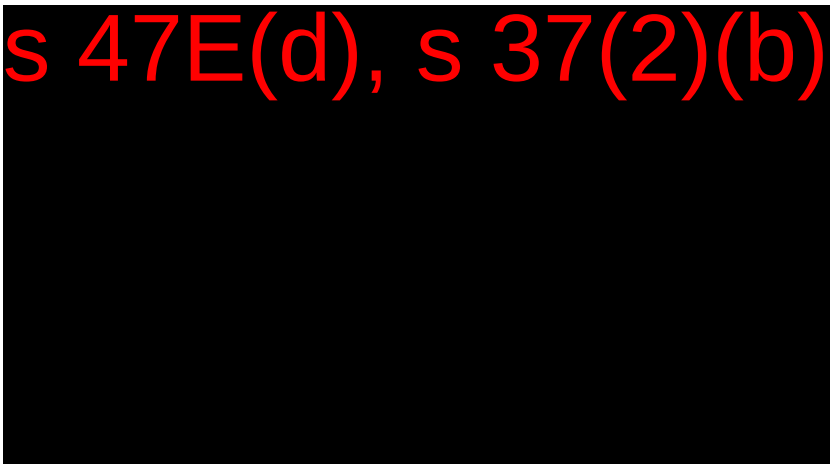
Timeline (Slide Layer)



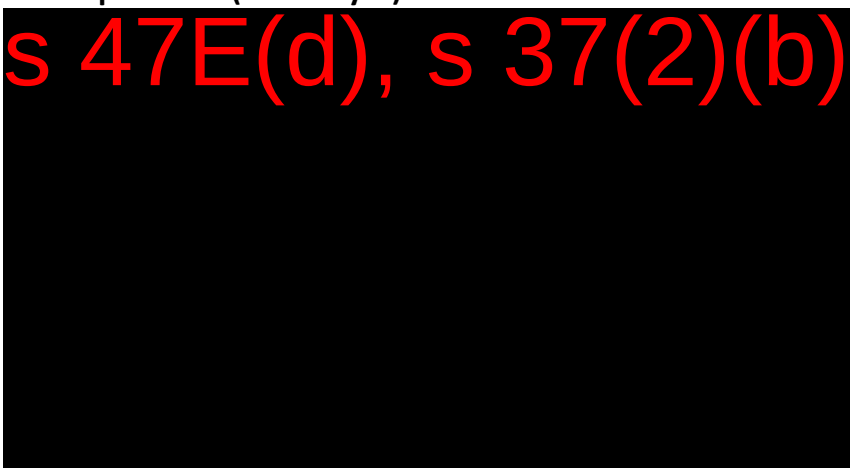
Map (Slide Layer)



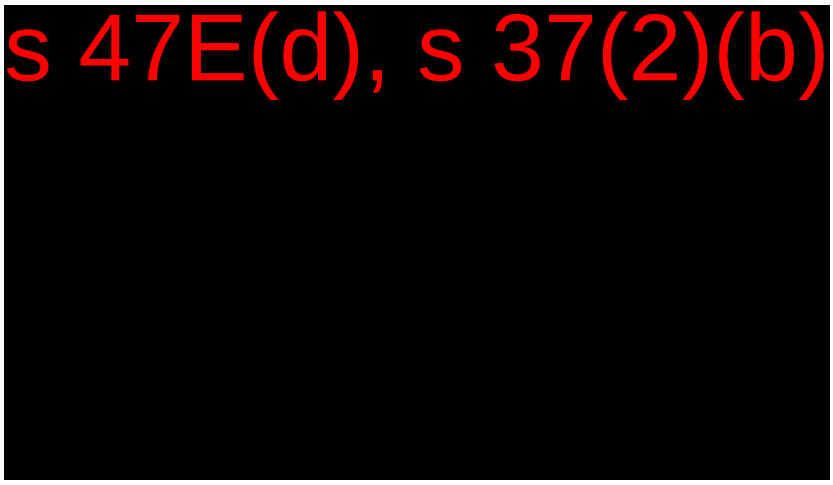
Browser Video (Slide Layer)



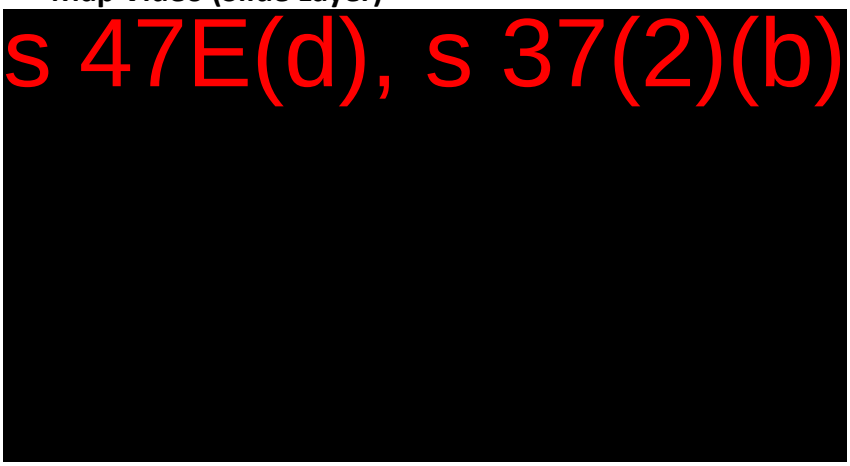
Graph Video (Slide Layer)



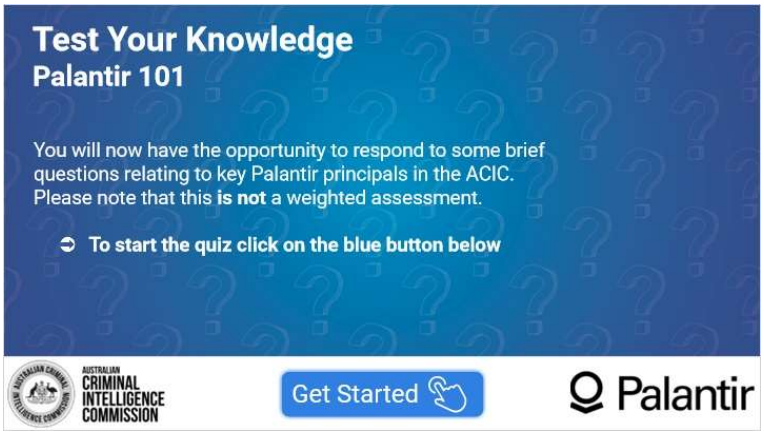
Timeline Video (Slide Layer)



Map Video (Slide Layer)



1.6 Test Your Knowledge - Palantir Gotham 101



1.7 How do we use Palantir?



Published by Articulate® Storyline www.articulate.com

Business Rules (Slide Layer)



Information Goverance (Slide Layer)



Access Governance (Slide Layer)



Integrity Assurance (Slide Layer)



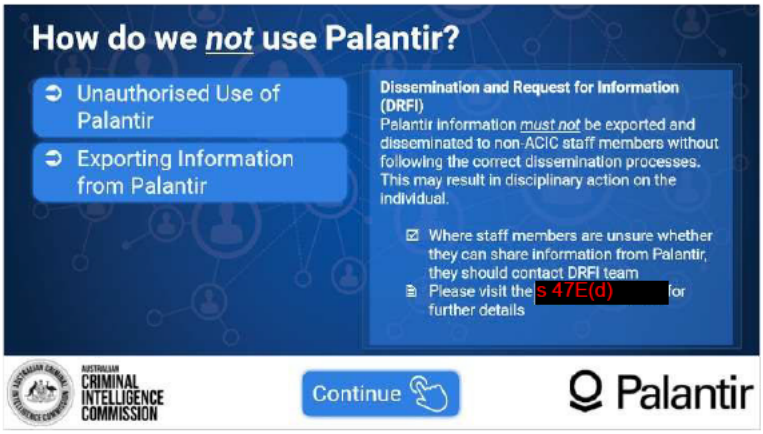
1.8 How do we not use Palantir?



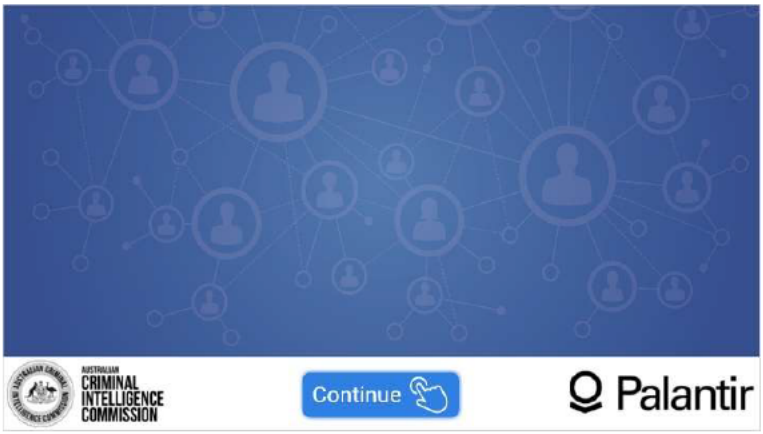
Unauthorised Use (Slide Layer)



Exporting Info (Slide Layer)



Case Study 2 (Slide Layer)

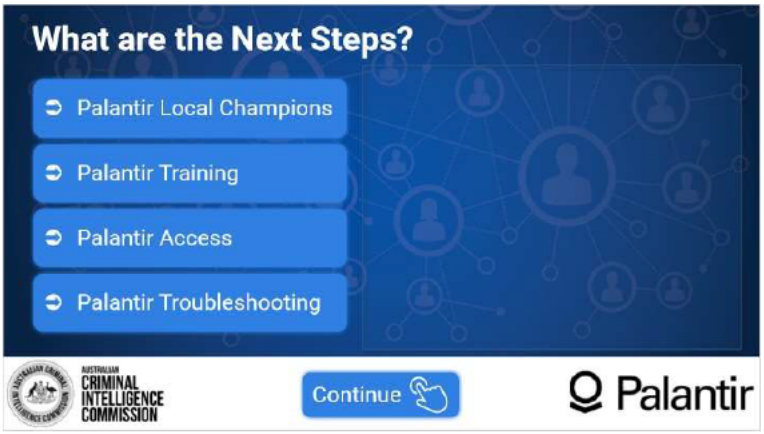


Palantir Ecosystem 2 (Slide Layer)

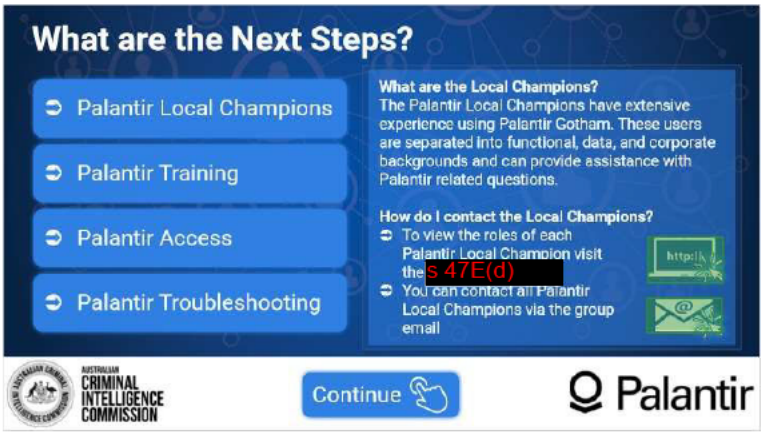


Published by Articulate® Storyline www.articulate.com

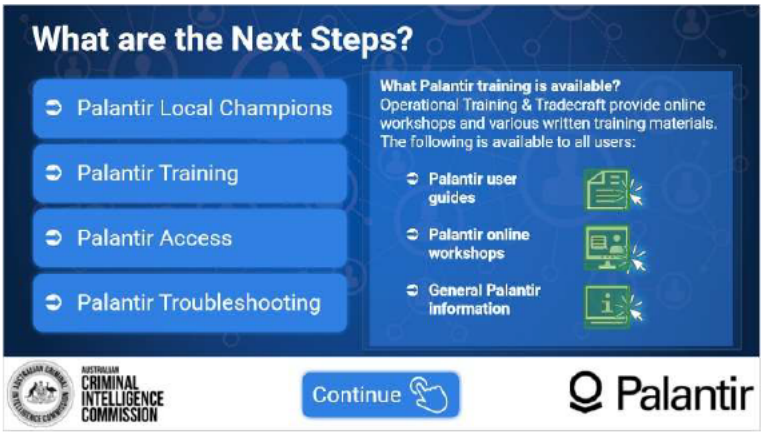
1.9 What are the Next Steps?



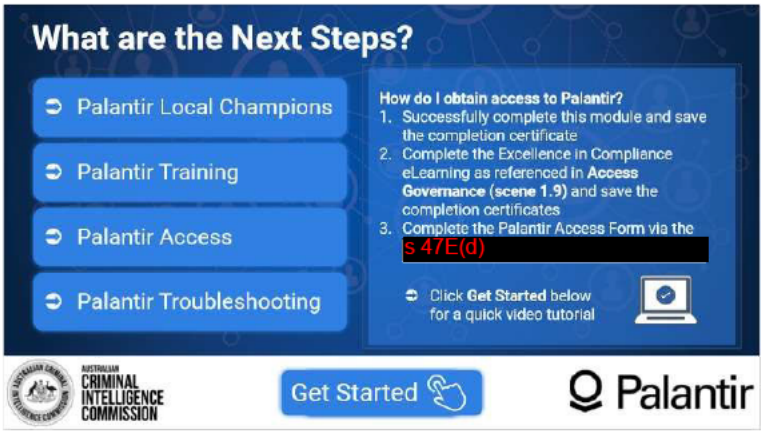
Local Champions (Slide Layer)



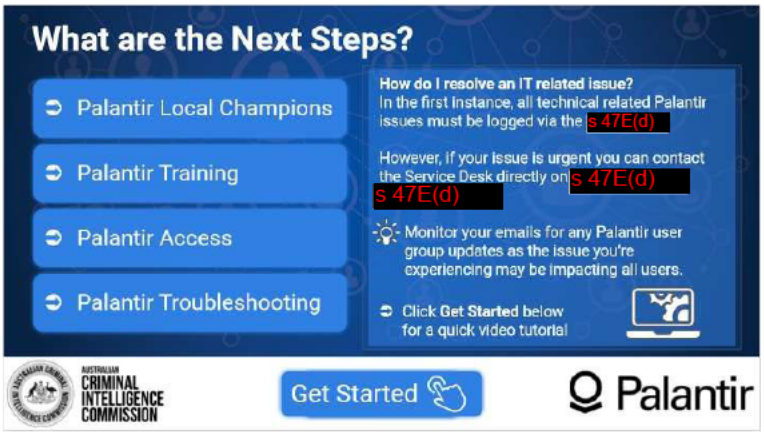
Training (Slide Layer)



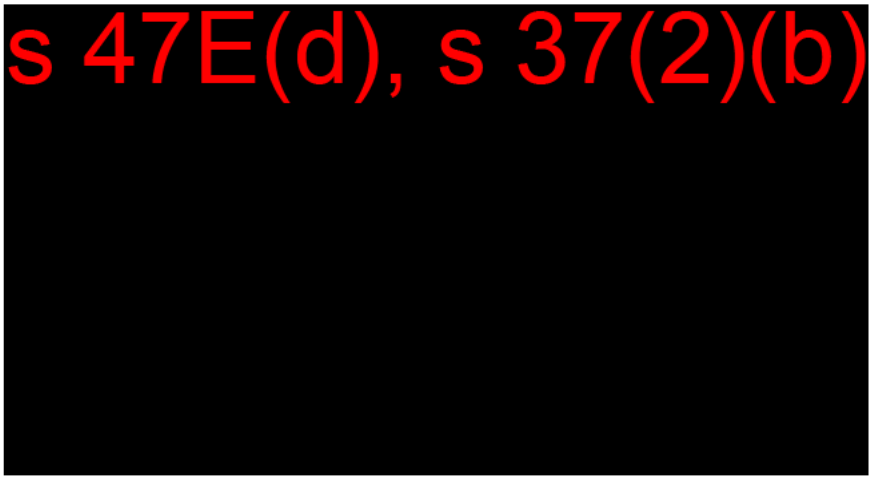
Access (Slide Layer)



Troubleshooting (Slide Layer)

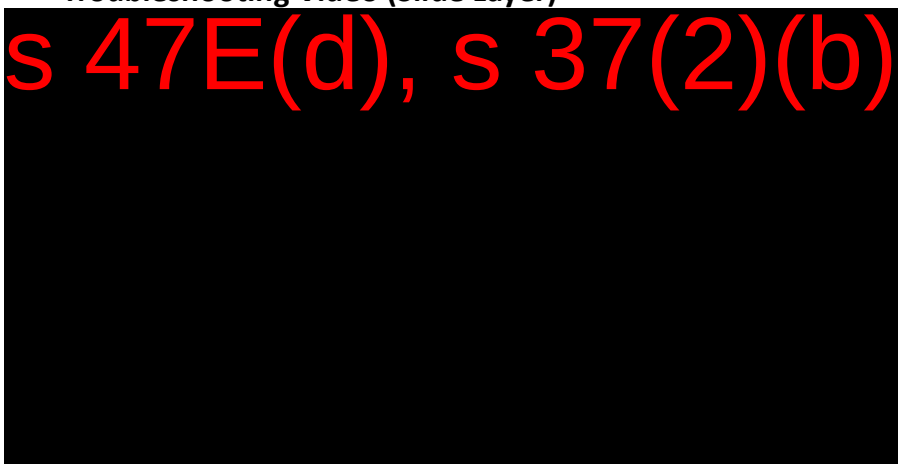


Access Video (Slide Layer)

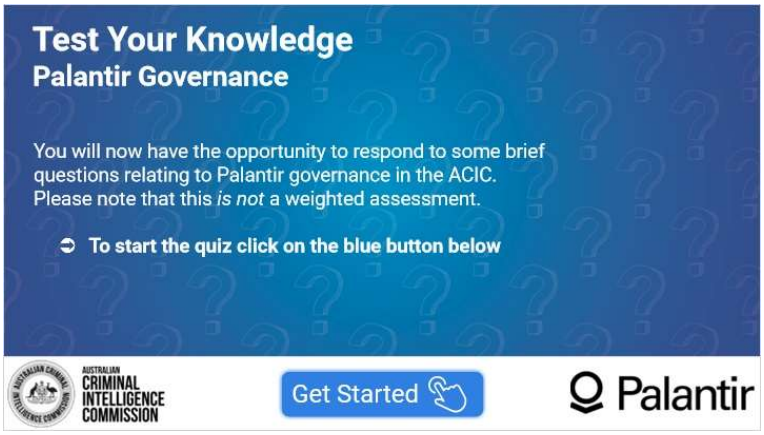


Published by Articulate® Storyline www.articulate.com

Troubleshooting Video (Slide Layer)



1.10 Test Your Kowledge - Palantir Governance



1.11 Need Assistance?

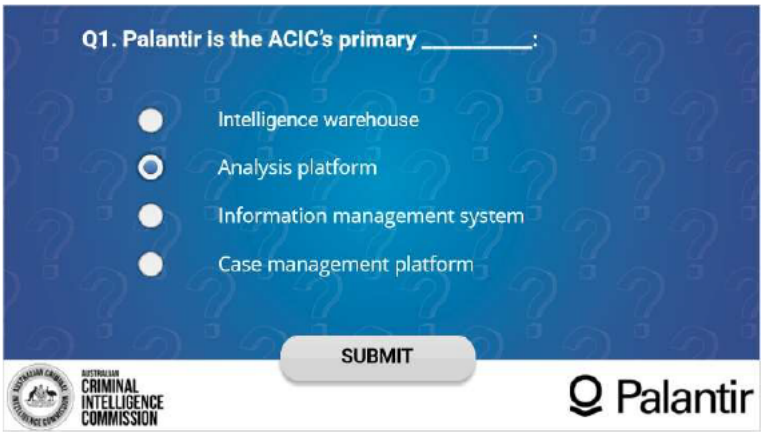


---

2. Test Your Knowledge - Palantir 101

2.1 Question 1

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
	Intelligence warehouse
X	Analysis platform
	Information management system
	Case management platform

Feedback when correct:

That is correct! Although Palantir stores information and intelligence, it's primary purpose is to operate as an analysis platform.

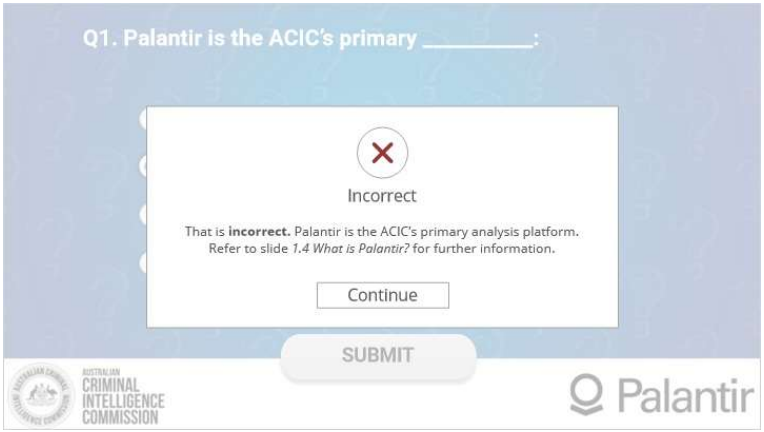
Feedback when incorrect:

That is incorrect. Palantir is the ACIC's primary analysis platform. Refer to slide 1.4 What is Palantir? for further information.

Correct (Slide Layer)



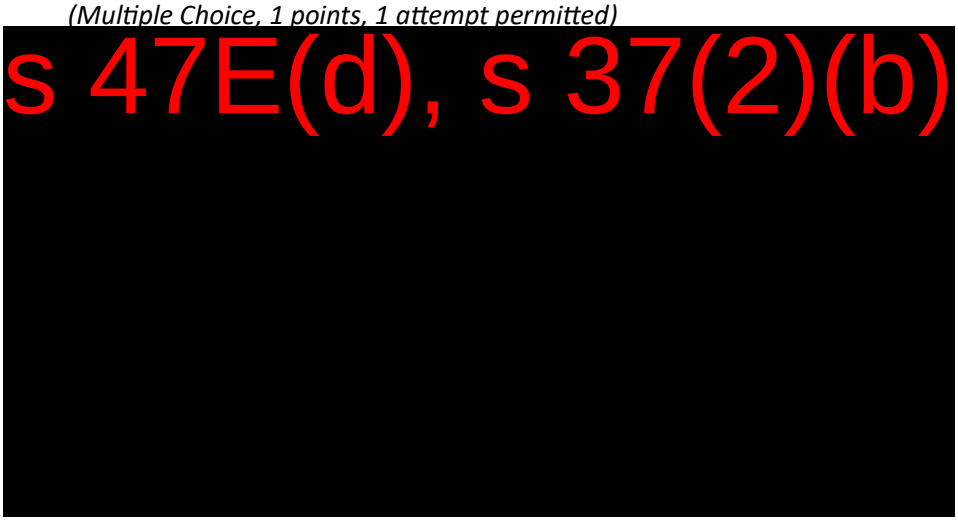
Incorrect (Slide Layer)

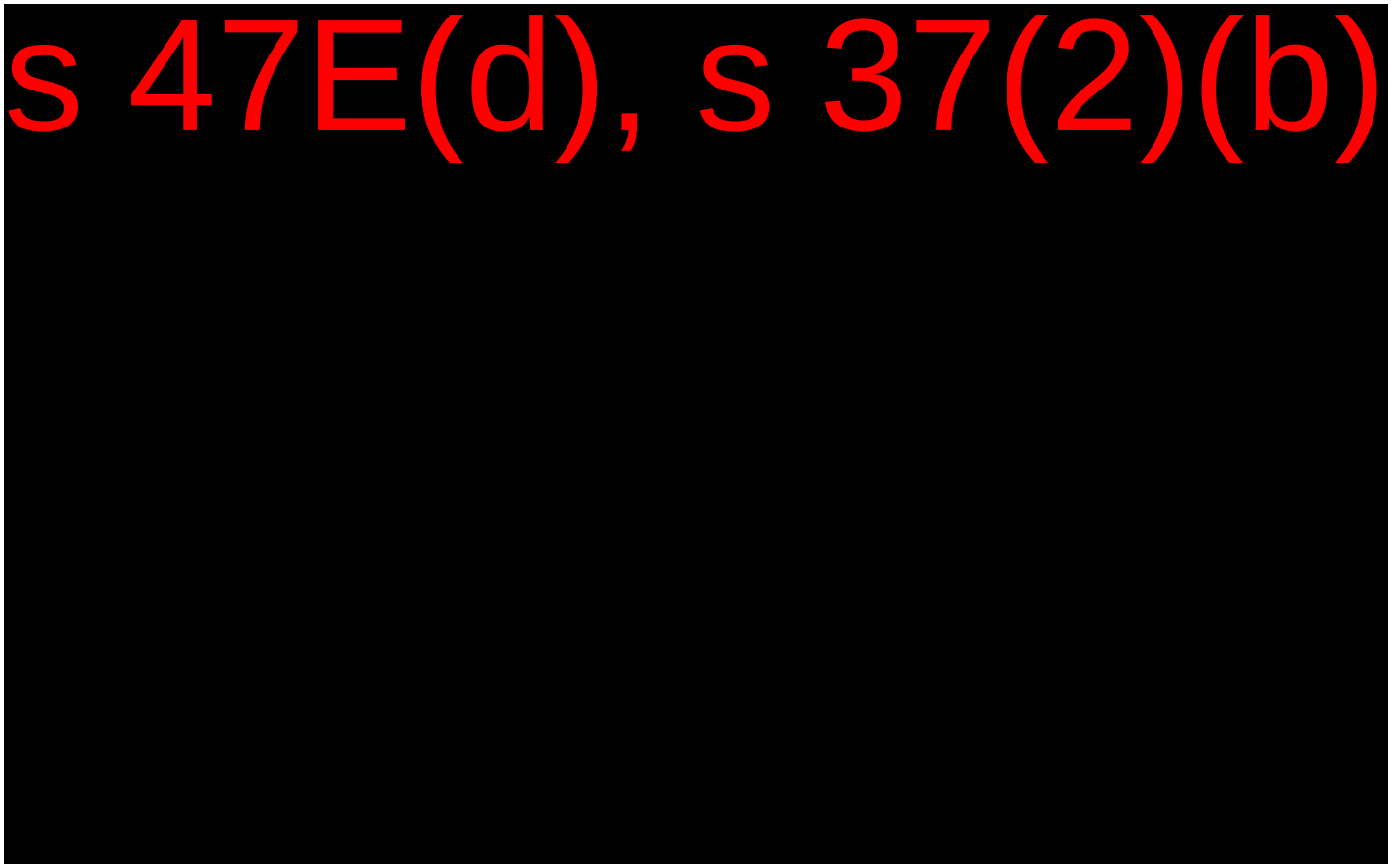


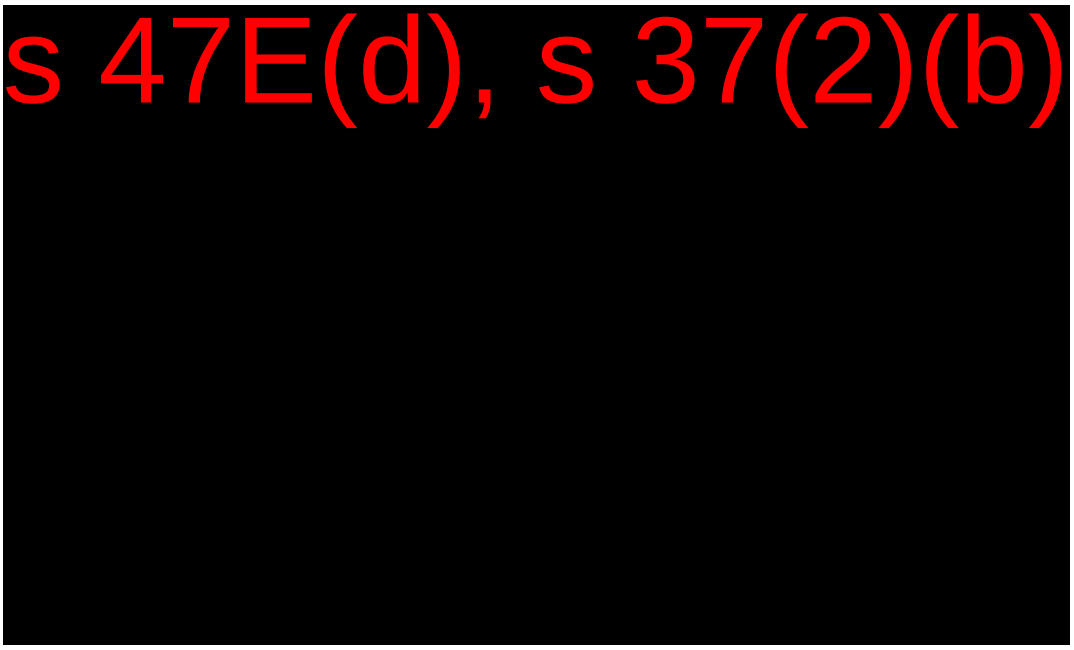
2.2 Question 2

(Multiple Choice, 1 points, 1 attempt permitted)

s 47E(d), s 37(2)(b)





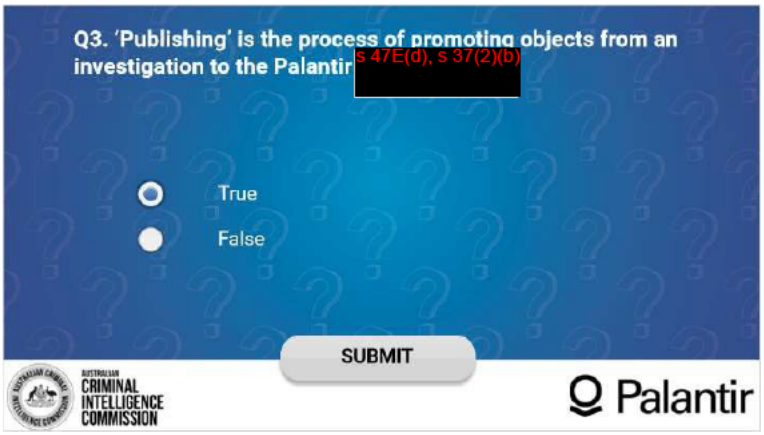


Published by Articulate® Storyline www.articulate.com

s 47E(d), s 37(2)(b)

2.3 Question 3

(True/False, 1 points, 1 attempt permitted)



Correct	Choice
X	True
	False

Feedback when correct:

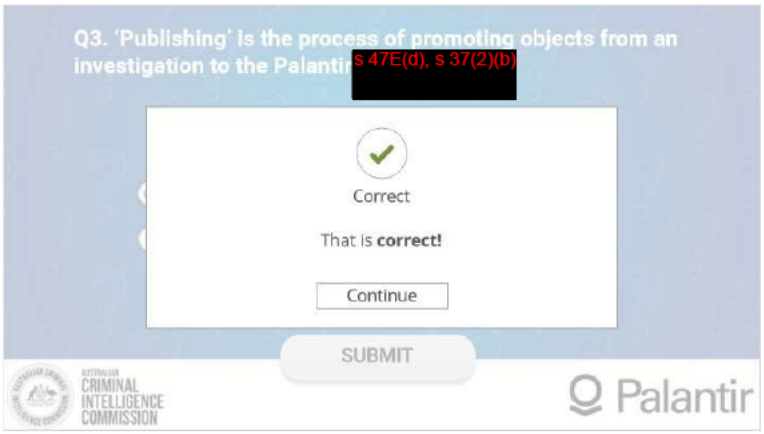
That is correct!

Feedback when incorrect:

Published by Articulate® Storyline www.articulate.com

That is incorrect. Publishing is the process used to promote objects from a users investigation to the Palantir s 47E(d), s 37(2)(b)

Correct (Slide Layer)

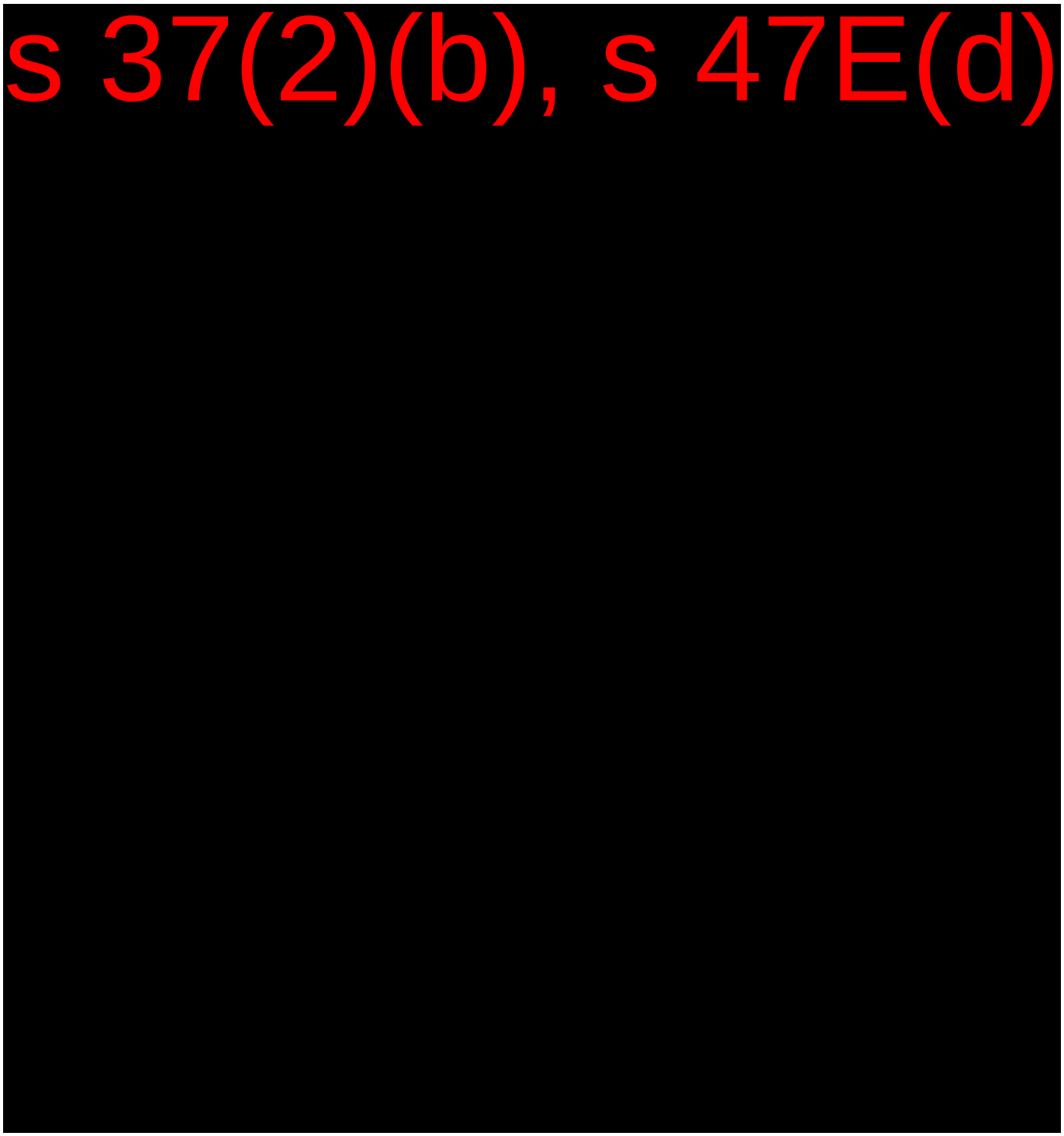


Incorrect (Slide Layer)



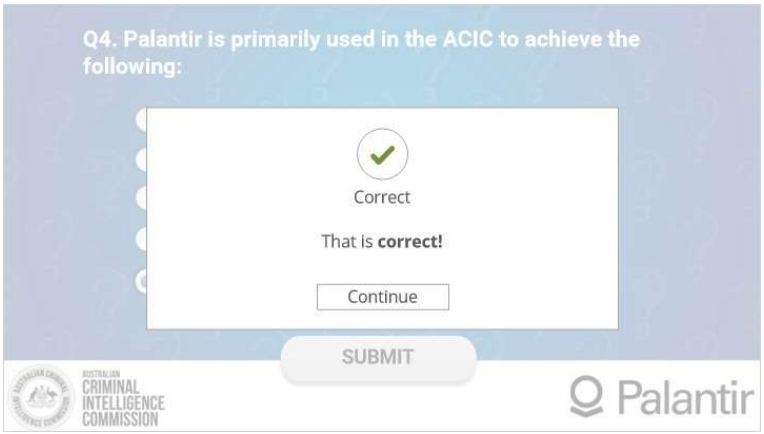
2.4 Question 4

(Multiple Choice, 1 points, 1 attempt permitted)

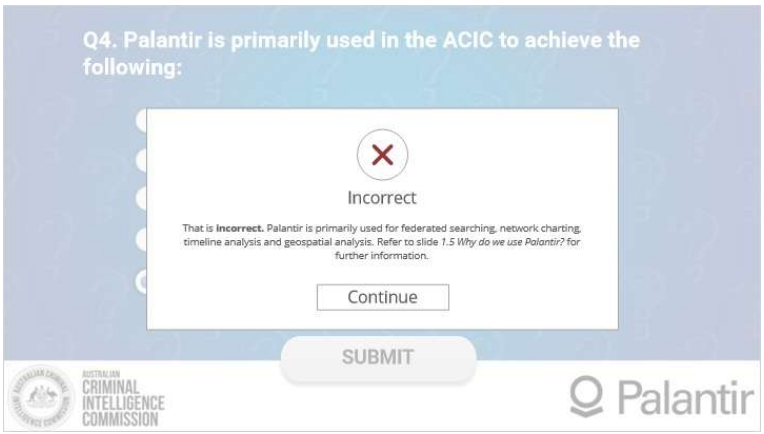


Published by Articulate® Storyline www.articulate.com

Correct (Slide Layer)

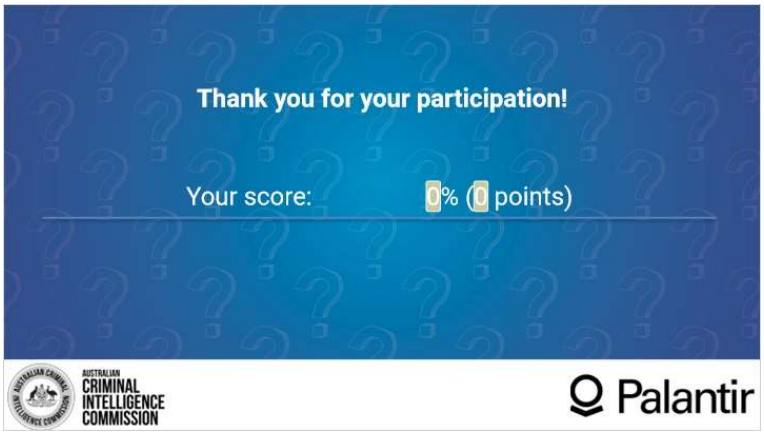


Incorrect (Slide Layer)



2.5 Results Slide

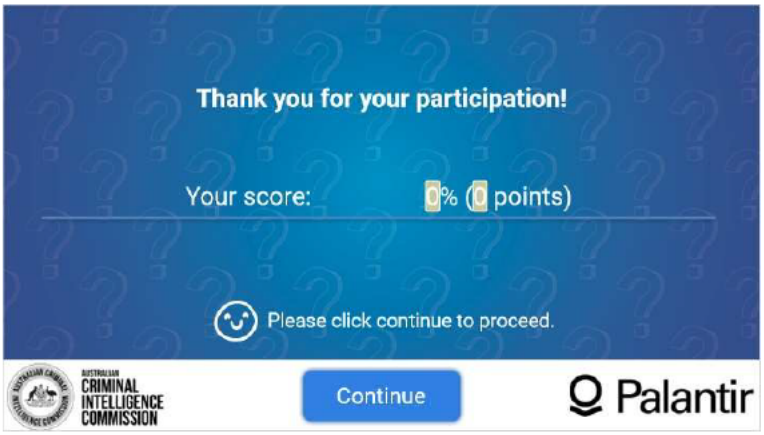
(Results Slide, 0 points, 1 attempt permitted)



Results for
2.1 Question 1
2.2 Question 2
2.3 Question 3
2.4 Question 4

Result slide properties	Passing Score
0%	

Continue (Slide Layer)



3. Test Your Knowledge - Palantir Governance

3.1 Question 1

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
	Outline procedures for automating document imports to Palantir
	Provide instructions for updating Palantir system settings
	Provide instructions for using Palantir applications and helpers
X	Outline the minimum standard for recording information/intelligence in Palantir

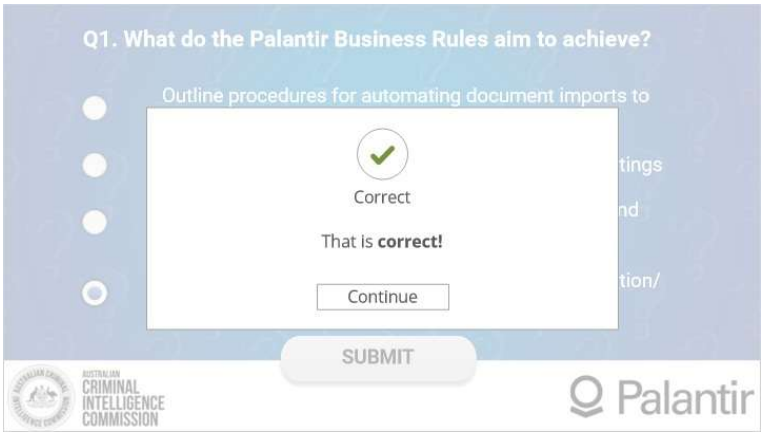
Feedback when correct:

That is correct!

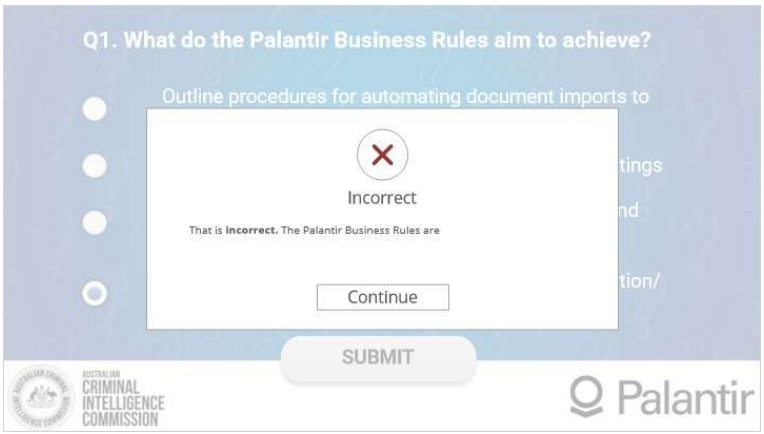
Feedback when incorrect:

That is incorrect. The Palantir Business Rules are procedures aimed to ensure a consistent minimum standard for recording criminal intelligence and information in Palantir.

Correct (Slide Layer)



Incorrect (Slide Layer)



3.2 Question 2

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
	An external partner asks via email or phone
X	Disseminated using the ACIC Disclosure Checklist
	When your manager asks you to forward information to them
	It is rated under the PROTECTED classification

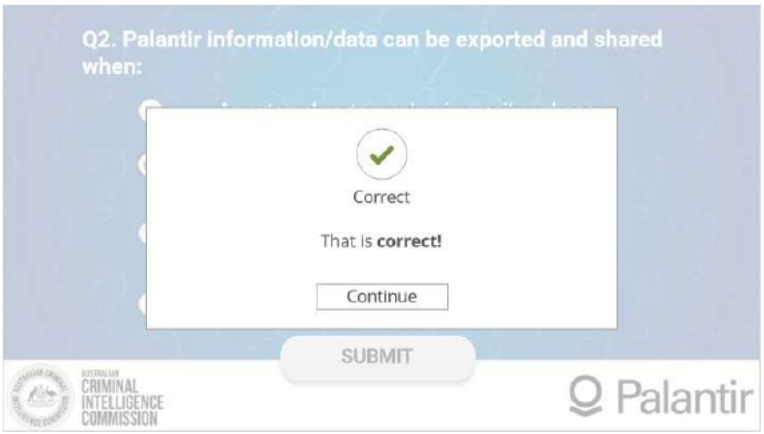
Feedback when correct:

That is correct!

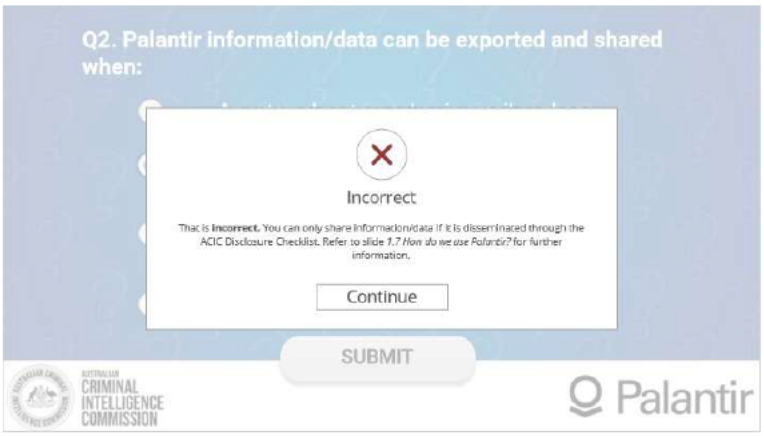
Feedback when incorrect:

That is incorrect. You can only share information/data if it is disseminated through the ACIC Disclosure Checklist. Refer to slide 1.7 How do we use Palantir? for further information.

Correct (Slide Layer)



Incorrect (Slide Layer)



3.3 Question 3

(True/False, 1 points, 1 attempt permitted)



Correct	Choice
X	True
	False

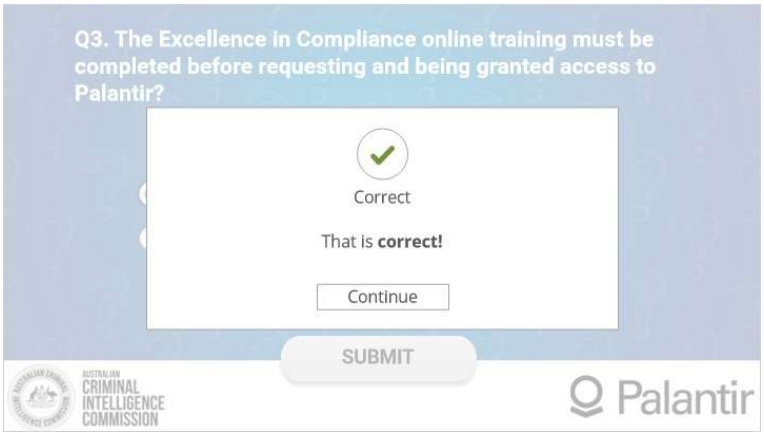
Feedback when correct:

That is correct!

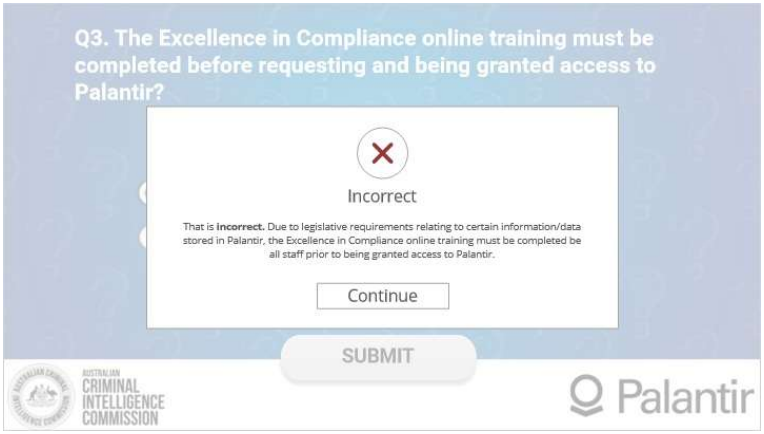
Feedback when incorrect:

That is incorrect. Due to legislative requirements relating to certain information/data stored in Palantir, the Excellence in Compliance online training must be completed be all staff prior to being granted access to Palantir.

Correct (Slide Layer)

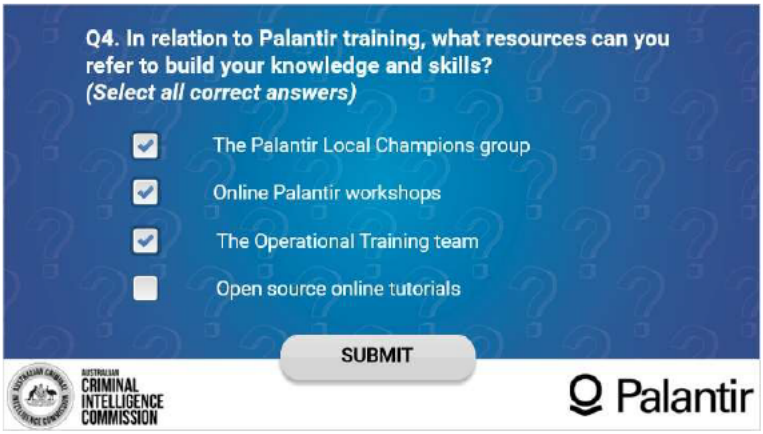


Incorrect (Slide Layer)



3.4 Question 4

(Multiple Response, 1 points, 1 attempt permitted)



Correct	Choice
X	The Palantir Local Champions group
X	Online Palantir workshops
X	The Operational Training team
	Open source online tutorials

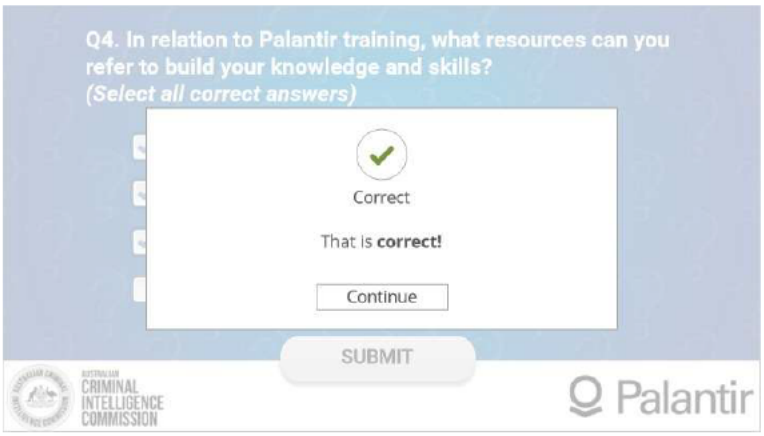
Feedback when correct:

That is correct!

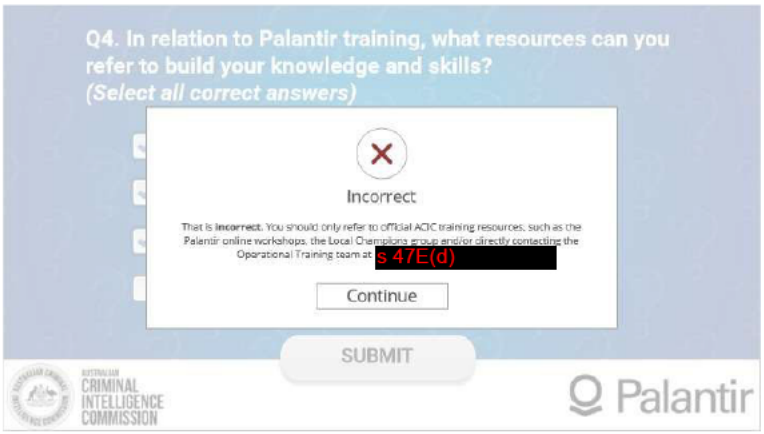
Feedback when incorrect:

That is incorrect. You should only refer to official ACIC training resources, such as the Palantir online workshops, the Local Champions group and/or directly contacting the Operational Training team at s 47E(d) [redacted]

Correct (Slide Layer)

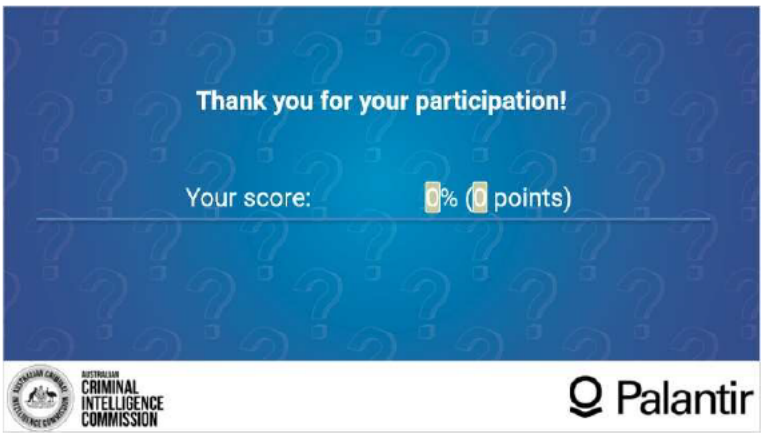


Incorrect (Slide Layer)



3.5 Results Slide

(Results Slide, 0 points, 1 attempt permitted)

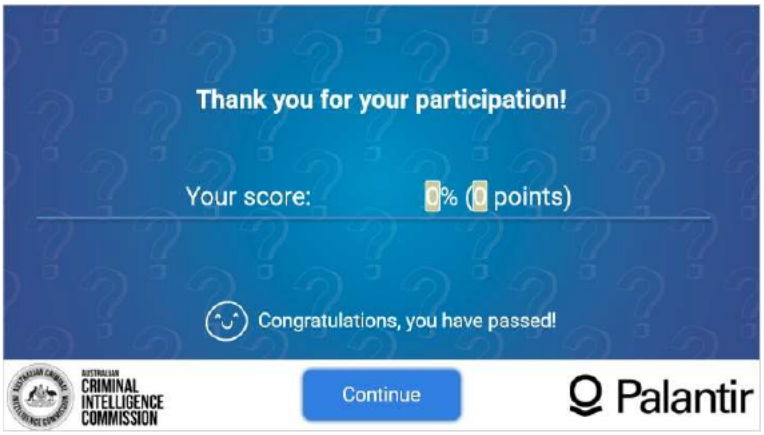


Published by Articulate® Storyline www.articulate.com

Results for
3.1 Question 1
3.2 Question 2
3.3 Question 3
3.4 Question 4

Result slide properties	Passing Score
0%	

Review (Slide Layer)



2. Quiz 1

Q2.1 Quiz 1 - Q1

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
	Palantir Foundry
	Palantir Riddler
X	Palantir Gotham
	Palantir Apollo

Feedback when correct:

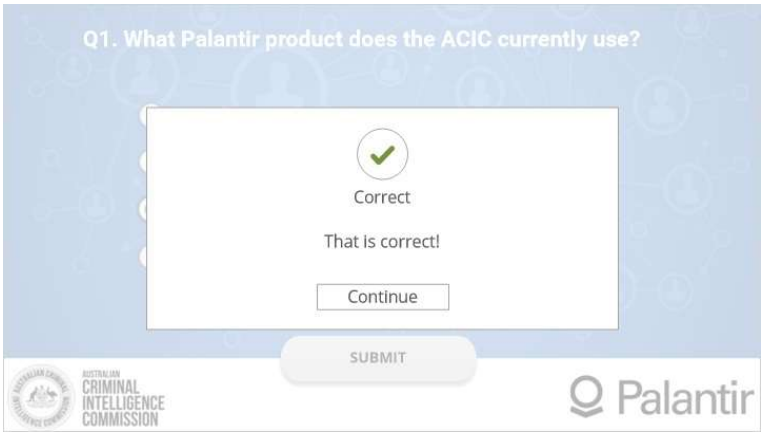
That is correct!

Feedback when incorrect:

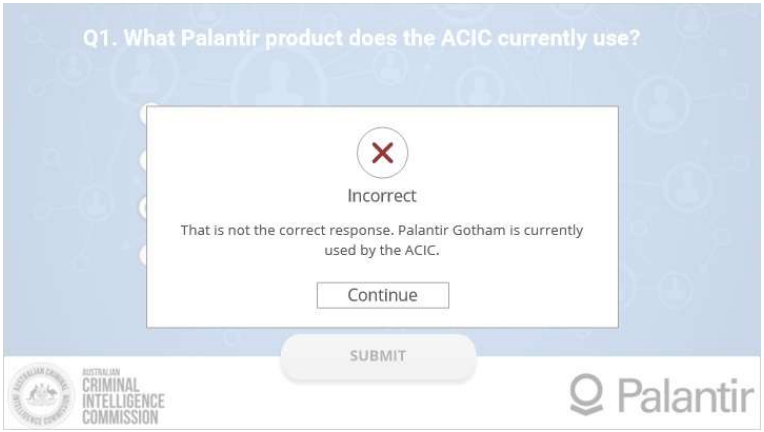
That is not the correct response. Palantir Gotham is currently used by the ACIC.

Published by Articulate® Storyline www.articulate.com

Correct (Slide Layer)



Incorrect (Slide Layer)



Q2.2 Quiz 1 - Q2

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
X	Integrated suite of analytical tools
	Complex criminal network
	Suite of data integration tools
	A feature of Palantir Foundry

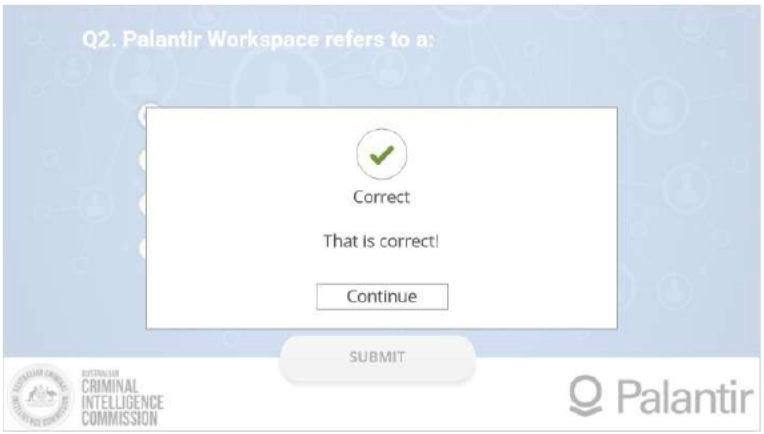
Feedback when correct:

That is correct!

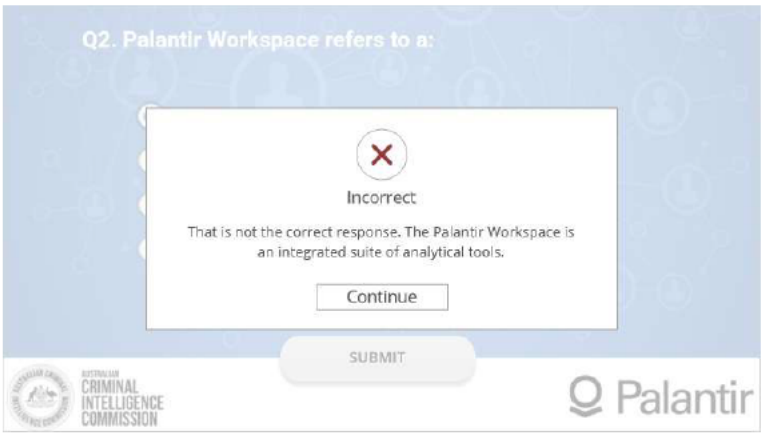
Feedback when incorrect:

That is not the correct response. The Palantir Workspace is an integrated suite of analytical tools.

Correct (Slide Layer)



Incorrect (Slide Layer)



Q2.3 Quiz 1 - Q3

(True/False, 1 points, 1 attempt permitted)



Correct	Choice
	True
X	False

Feedback when correct:

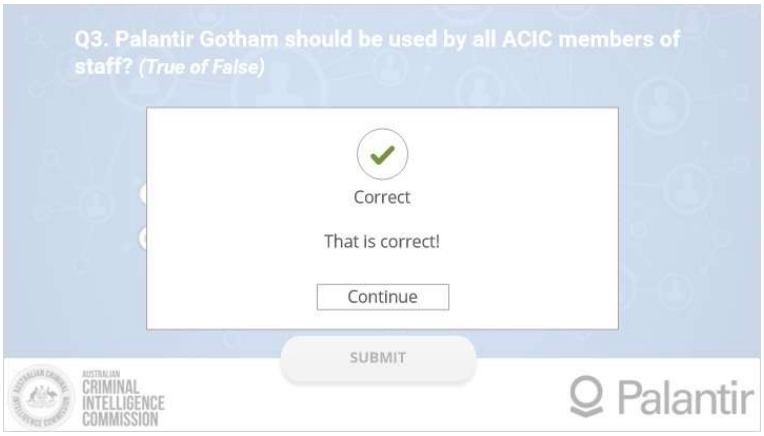
That is correct!

Feedback when incorrect:

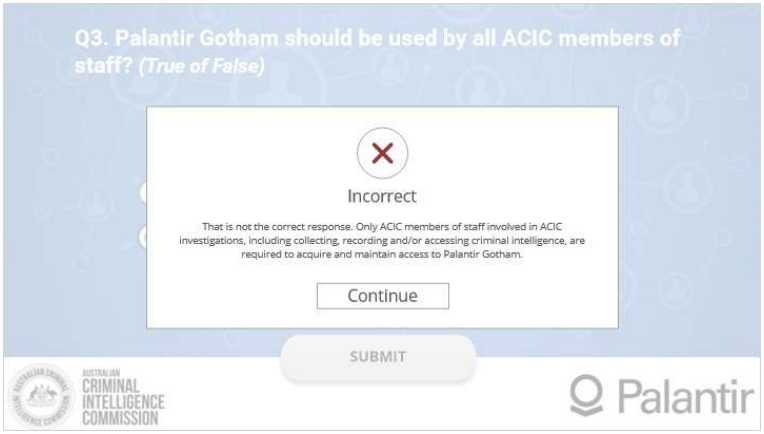
Published by Articulate® Storyline www.articulate.com

That is not the correct response. Only ACIC members of staff involved in ACIC investigations, including collecting, recording and/or accessing criminal intelligence, are required to acquire and maintain access to Palantir Gotham.

Correct (Slide Layer)

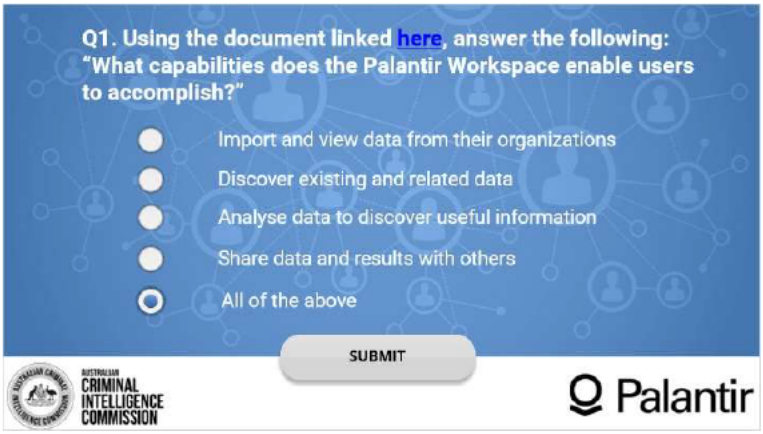


Incorrect (Slide Layer)



Q2.4 Quiz 1 - Q4

(Multiple Choice, 1 points, 1 attempt permitted)



Correct	Choice
	Import and view data from their organizations
	Discover existing and related data
	Analyse data to discover useful information
	Share data and results with others
X	All of the above

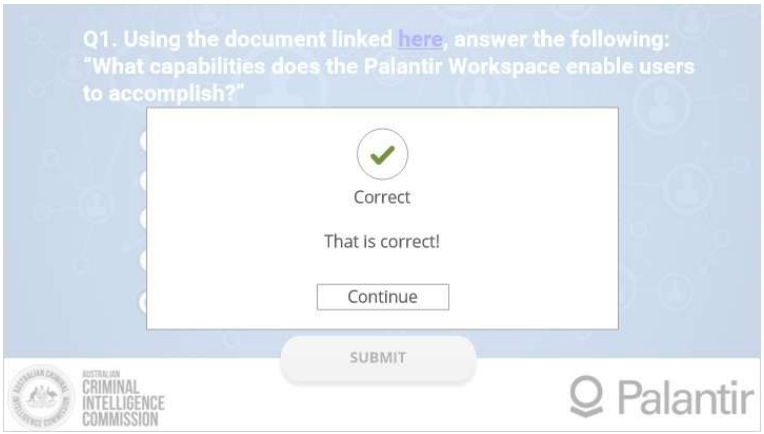
Feedback when correct:

That is correct!

Feedback when incorrect:

That is not the correct response. The Palantir Workspace enables the ACIC to import and view data, discover existing and related data, analyse data to discover useful information, and share data and results with others.

Correct (Slide Layer)



Incorrect (Slide Layer)

