CHAPTER

Causal Map App

CONTENTS

Contents

Getting Started

🎉 Welcome to CausalMap!!

📺 If you like learning through video, there are some short Causal Map videos to get you started: here. Or keep reading this Guide!

Try this: code the short example project:

- We've created a practice project for you called `example-short-[your-username]`. This gives you your own copy of our tutorial project to experiment with freely. **You should see it** in the Project Dropdown Menu at top left of the app.
- **Select source** `1` in the source selector. This will display the text from this source. This source is already partially coded for you.
- **Practice highlighting causal claims** in the Source Text Viewer to **create links** in the Create Links tab.
- **View, edit or delete existing links** in the Source Text Viewer in the Create Links tab by clicking on the highlighted sections of text or on the corresponding links in the Map.
- **Watch** the causal relationships grow in the Map panel

Try this: view some examples of what you can do.

These examples are views of a real-life, anonymised QuIP project in the field of international development.

- Main factors map.
- Main factors table.
- Consequences of increased knowledge map.

You can find them in the Project Dropdown Menu.

Get help / read the documentation

The documentation you are reading now is embedded in the app via the Help System and is also available as a standalone Guide.

Each section corresponds to a different part of the interface.

Within the app, you can:

- click the blue help buttons to find out more about the different parts of the app
- click the `[Help]`(#help-system) button at top right to open and search the help documentation.

But what even is causal mapping?

This Guide is all about how to use the app. For example it tells you how to use the different filters. If you want to find out more about the ideas behind those filters, we have a new "Ideas Garden"

with a more discursive look at the theory and everything that surrounds causal mapping. It's a work in progress but we are adding material every day, please bear with us.

Someone shared their work with you?

If they have to a particular map (or table), click the link and you will be automatically taken to view the corresponding map (or table), providing they have given you the correct permissions.

Alternatively if they have just asked you to log in at the app to explore a file name, say "project-x", log in and click on the dropdown list on the left-hand side and click on `project-x` to load it. If you can't find that file name, it means they haven't correctly shared it with the email address you logged in with.

How CausalMap Works

1. **Create Projects**: Create a project and upload PDF or docx documents into it as source material for analysis (we call each document a "source"). You can have multiple projects, and each project usually contains multiple sources.
2. See Projects Panel and Upload Documents.
3. **Create Causal Links manually**: Identify and highlight causal claims within the source text; each highlight becomes a new causal link. After you have created multiple links you can view them as a causal map.
4. Use the Create Links tab and view results in the Map panel.
5. **Create Causal Links using AI**
6. See AI Coding.
7. **Filter & Analyze**: Apply filters to your causal links focus on specific aspects of your data
8. Use the Analysis Filters: Do qualitative causal analyseis on the selected links by filtering manipulate those links.
9. Explore Map, Factors, Links, and Statistics.
10. **Collaborate**: Work together live with multiple users
11. Manage sharing in Sharing and Permissions.

New features

The previous version of Causal Map, version 3, was already, as far as we know, the only software dedicated to causally coding causal claims within texts. Version 4 improves over Causal Map 3 in the following ways:

Speed and stability

- Fast loading and editing
- Resilient to poor internet connectivity
- Scalable to hundreds or thousands of concurrent users

Uploading and organising data

- Simplified data model to make it easier for you to import and manage your sources: we no longer break down source texts into separate statements. We treat each source text as one entire text. In cm4, there is no such thing as a statement.
- Easily create "Custom Columns" for each source (such as gender, location etc)
- Edit the data for your "Custom Columns" with a spreadsheet-like interface
- Complete management of all of your projects
- Simple upload of PDF or DOCX documents
- Tags to help you organise multiple projects
- Note that in cm3 we used to call projects "files"
- You can upload a project exported from cm3 as a new cm4 project

Managing labels

- New, more powerful and easier to use bulk editing of labels in the links table and the factors table.

Filtering and analysing

- Almost all the existing links filters from cm3 are available plus
- the option to include multiple versions of the same links filter, e.g. to narrow down a selection of links by different criteria successively
- optional semantic filters like `cluster` and `soft recode` **Require an AI subscription**
- Analyse data with new pivot tables and graphs

Sharing and collaborating

- Anonymous login option so that anyone can view your work without logging in
- URL-based state saving for bookmarking (the same URL always takes you back to the same view)

These features require a Team subscription

- Real-time collaboration: Live updates when collaborators add links
- Interactive maps for live demonstrations

Help system

- Built-in help system
- Help drawer with links to each section
- Same contents used for separate Guide with links to each section

AI Coding

- Optional AI-powered state-of-the-art, paragraph by paragraph coding assistance. We call this "Human first, AI next".

Causal Map 3 features which will probably *not* be implemented:

- Deep support for standard questions across multiple sources
- Special treatment for closed questions
- Ability to view the text of multiple sources at once.

The left and right panels

The app uses a two-pane layout with a draggable border between them (default split 30:70).

> The left hand side of the app is all about selecting sources then creating and filtering links.
>
> The right hand side (the pink tabs) is all about presenting the results.

Left-hand side

- Project Dropdown Menu: select a project including its links and documents
- Sources Dropdown Menu: choose one or more sources. **(Leaving it empty includes *all* sources).**
- Create Links tab: Read and code the text of the selected source(s). If multiple sources are selected, the first is shown.
- Filter Links tab: Do qualitative causal analyseis on the selected links by filtering and manipulating them.

The Links Pipeline

The diagram shows the Links Pipeline: The top four boxes here correspond to the left side of the app and are called the "Links Pipeline": each step selects and filters links. The resulting links are then displayed in the pink output tabs on the right side.

```
graph TD
    A["📁 Select Project"] --> B["📄 Select Sources&#10;(Documents)"]
    B --> C["🔗 Extract Links&#10;(Causal relationships)"]
    C --> D["🔍 Apply Analysis Filters&#10;(Factor labels, paths, etc.)"]
    D --> E["📊 Display Results"]

    E --> F["🗺️ Map&#10;(Network visualization)"]
    E --> G["🏷️ Factors Table&#10;(Causes & effects)"]
    E --> H["🔗 Links Table&#10;(Relationships)"]
    E --> I["📈 Pivot Tables&#10;(Charts & analysis)"]

    style A fill:#e1f5fe
    style B fill:#f3e5f5
    style C fill:#fff3e0
    style D fill:#e8f5e8
    style E fill:#fce4ec
    style F fill:#fff9c4
    style G fill:#fff9c4
    style H fill:#fff9c4
    style I fill:#fff9c4
```


Right-hand side (pink tabs)

- Outputs: these all show the same filtered links from the Links Pipeline but in different formats
- Map: visual network of links
- Factors: editable factor list (toggle available to bypass analysis filters)
- Links: editable links table (toggle available to bypass analysis filters)
- Pivot Tables: additional analysis and charts

Right-hand side (other tabs)

The right-hand side also contains other tabs not influenced by the pipeline:

- Help: help drawer and docs
- Projects: manage projects
- Sources: manage sources
- Settings: application preferences
- Account: your account
- Logs: application logs
- Bookmarks: saved views
- Responses: AI logs and usage

Tips for using the app

Tips for Using the Dropdown Menus

There are many dropdown menus throughout the app.

- Most dropdowns allow multiple selections: you can select more than one thing at once
- Most dropdowns allow you to type and create new entries which are not already in the list.
- Type part of a word and click "Create new..." to add new items
- Press Enter to complete selections
- Pressing Tab always moves you to the next field (doesn't complete selection). See also Search/replace for bulk editing patterns.

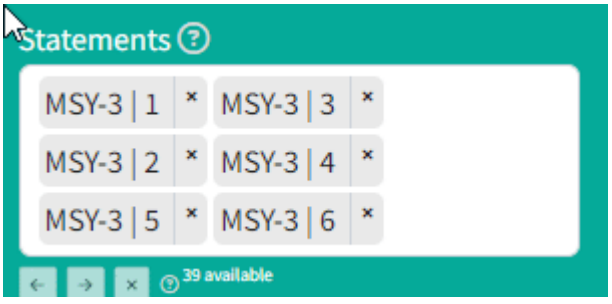
Edit an item

- **Backspace editing**: Position cursor after an existing selection and press backspace to edit it

Click to the right of the item, then press backspace to edit the label, and press tab to complete.

Delete the first and subsequent items using the keyboard

Click after the last item, use the left arrow on your keyboard to go back to previous items, then press Delete on your keyboard to delete items after the cursor.



Tips for Using Tables

Most tables include:

- **Checkboxes** ☑ for selecting multiple rows
- **Bulk action buttons** when you have selected one or more rows (edit, delete)
- **Action buttons** within individual rows to apply actions (edit, delete etc) just to that row
- **Sorting** by clicking column headers
- **Filtering** using the filter row below headers

- **Pagination** with 10/25/50/100 items per page
- **Re-ordering columns** by dragging the column headers

Tips for Using Prompts and other text windows

Require an AI subscription

When you use text windows, your texts are automatically saved so you can reuse them later.

Text history is available in:

- **AI Coding** (Process Sources tab)
- **AI Answers** (Answers tab)
- **Map Vignettes** (Vignettes tab)
- **Soft Recode filters** (both label prompts and magnet lists)
- **Auto Recode filter**

How to use text history:

- **Dropdown menu**: Shows your previous prompts with project name and date/time
- **< and > buttons**: Navigate between newer and older prompts
- **Text area**: Shows the selected prompt and lets you edit it
- **Expand button**: Optionally dit your text in a larger, more convenient text editor with multiple cursors, search/replace etc.
- **Trigger button**: Runs the AI with your current prompt and saves it to history

How prompts are organized:

- Your current project's prompts appear first (most recent at top)
- Then prompts from other projects you can access
- Each prompt shows when it was last used
- Duplicate prompts are automatically removed

Using the controls:

- Select any prompt from the dropdown to load it
- Use < and > buttons to move through your text history
- Edit the prompt in the text area as needed
- Click the action button (Process, Ask, Generate, etc.) to run it

Your most recent prompt automatically appears when you open each AI feature.

Tip: Any lines beginning with // in your prompt will be recorded in the history etc but not actually sent to the AI. You can use this to make notes e.g. at the top of your prompt: "//Sarah's version with tweaked summary"

Projects Bar

🗂️ **What you can do here:** Choose which Project (project) you want to work on. Use the File menu for quick actions like creating new Projects, uploading documents, or sharing your work with others.

- A small locked indicator at the top-right of `#project-selector-header` shows only when a project is read-only.
- An archived icon appears only when the project is archived. Archived projects are automatically read-only and only visible to owners, collaborators, and admins (not visible to the public even if marked public).

File menu

Quick access to common actions:

Manage the Current Project

- **Edit** : Modify settings and sharing
- **Upload sources** : Add documents to the current Project
- **Clone** : Create a complete copy of the Project under a new name
- **Clone filtered** : Create a copy of the current Project but only containing the sources and links as currently filtered.
- **Archive** : Hide from main list
- **Delete all links** : Remove every causal link in this Project (sources remain)
- **Download** : Export as XLSX
- **Versions** : Restore and create backups of this Project
- **Manage projects** : Opens the Projects tab as a shortcut

Link to the Current Project

- **Copy link** : Get short bookmark URL (e.g., `?bookmark=abc123`) to current app state
- **Copy formatted link** : Get HTML link with bookmark ID as text (e.g., `<a href="...?bookmark=abc123">#abc123</a>`) for documents and emails

Manage projects

- **New Project** : Create an empty Project which you can then import sources into
- **Import XLSX** : Import a complete new Project from Excel (e.g. another Causal Map 4 project which has been exported as XLSX). You can use this for "round-tripping": downloading, conducting some bulk operations or other tweaks within Excel and then uploading it again.
- **Update Sources in current project** : Upload Excel sheet with just a sources tab in standard format with updated sources data e.g. additional or corrected custom columns. You can use this for "round-tripping": downloading, conducting some bulk operations or other tweaks within Excel and then uploading it again.
- **Import cm3** : Import a complete new Project downloaded from CausalMap3

Project Dropdown

- Lists all the Projects you have created or been invited to, from which you open just one
- After changing the project, the rest of the app **resets to defaults**: sources filter, all links filter pipeline filters, and deck filter (shows all bookmarks)
- On startup, the app auto-selects the most recent **viewable** project (owned by you, shared with you, or public).
- Admins can see all projects in the Projects table, but the dropdown never auto-selects or loads a non-viewable project; admins may open the Edit Project modal for non-viewable projects from the table only.

Project Details button

- A small pencil button sits to the right of the Project Dropdown.
- Clicking it opens the Project Details screen, the same as you get by clicking the first item in the File menu, which we describe next:

Project Details screen

Manage every aspect of the current project. You can reach this management pane by clicking Edit from the File menu. Manage other projects by clicking the edit button in the corresponding row of the Projects Panel.

📄 **What you can do here:** - **Open it**: File → Edit, or the edit icon in the Projects table. - **When it appears**: Also opens on project load/change unless you turn it off for this project. - **Save vs instant changes**: - Use the *Save* button to apply changes to **Name**, **Tags**, and **Description**. - **Archived**, **Locked**, **Public**, and **Collaborators** update instantly.
Details

- **Name**: Rename the project. Click *Rename* to save.
- **Tags**: Comma-separated tags for quick grouping/searching.
- **Description**: Free-text notes about the project.
- **Edit codebook**: Toggle to reveal a text area where you can list factor labels (one per line). These are added to the cause/effect dropdowns in the link editor; existing options are kept.
- **AI Processing Region**: Choose where AI processing occurs for GDPR/data residency compliance:
- **EU (Belgium - europe-west1)** - Default. Recommended for EU data residency requirements.
- **UK (London - europe-west2)** - UK has GDPR adequacy decision, suitable for EU/UK compliance.
- **US (Virginia - us-east5)** - US East region.
- Setting is saved per-project and auto-saves on change (with confirmation warning).
- All subsequent AI coding for this project uses the selected region.
- **Archived**: Hide the project from the main list and make it read-only. Archived projects are only visible to owners, collaborators, and admins (hidden from public view even if marked

public). Applied immediately. Unarchiving restores normal visibility and editability.
- **Info line**: Created / Modified / Owner, plus counts for links/sources/words, and quick actions:
- *Versions*: Open the versions manager.
- *Delete embeddings*: Remove factor embeddings for this project (advanced).
- **Show on open**: Toggle "Show this screen when opening this project" at the top to auto-open or suppress this screen for this project.

Sharing

- **Locked**: Make the project read-only. Editing is disabled until unlocked. Applied immediately.
- **Public**: Allow all signed-in users to view the project (read-only). Applied immediately.
- **Collaborators**: See current collaborators, add by email, and choose permission:
- *Viewer*: Read-only
- *Editor*: Read & write

Bookmarks

- If available, view saved “bookmarked” views of your data and open the Bookmark Manager.

This screen also shows when a project loads or is changed, except:

- for new users (the help drawer is still being opened to welcome them),
- if you have already clicked "don't show" for this project.

Sources Bar

📄 **What you can do here:** Choose which source documents (e.g. interviews or reports) from your current project you want to focus on. You can select one or more sources. Use this to narrow your analysis to specific interviews, reports, or other source materials. - The text of the selected source is shown below in the Create Links panel. - Selecting these sources also fetches only their links and no others, starting off the [Links Pipeline](../links-pipeline/): only the links from the currently selected sources are available for further filtering, and are finally shown in the output tabs.

Sources Dropdown

- Contains IDs of all sources in current project
- Select one or more sources
- Search by typing

- Ordering: alphabetically by source title (with any leading "Source " stripped), falling back to ID when title is missing. Next/Previous navigation uses the same order. - Opening behavior: if a source is currently selected, opening the dropdown will start at the next source after the current one (wrap-around at end). If no source is selected (empty dropdown), it starts at the first source. - The dropdown does not auto-open when sources are updated; it only opens on explicit user interaction (click/focus). -->

Source Groups sub-panel{#source-groups-sub-panel}

What this does: Filter your analysis by participant demographics or document characteristics, using the [custom columns](../custom-columns/) you have defined for your project. For example, show only responses from "women aged 25-35" or interviews from "urban areas." Perfect for comparing how different groups see causal relationships.
Controls:

- a pre-populated dropdown called **Field** listing metadata fields plus title and projectname
- a multi-select **Value** dropdown (filtered by Field)
- optionally sampling buttons for deterministic subsets:
- **Random 5** – loads five random sources from the whole project (also, Random 10, Random 20 etc)
- **Random 5/Group** – after choosing a Field, loads up to five random sources *for each value of that Field*.
- **Clear** button

The sampling buttons ake a random selection but in such a way that the same sources will be chosen if you click the same button again.

The second, "Value" dropdown is filtered to show only valid values for the selected field. Previous/next buttons cycle through values of the selected group.

The effect is to retain only links where the selected custom column has the selected values.

This dropdown is NOT a filter and it does NOT get saved/restored in URL. It is a loader: when we click it, the app automatically loads corresponding sources into the sources selector. These sources then DO form part of the links pipeline and ARE restored from the URL.

There is a similar filter in the Analysis Filters.

Create Links tab

Qualitative causal mapping involves taking passages of text, e.g. from interviews or documents, and identifying sections which make causal claims. We highlight each of these sections and specify a causal factor at each end of each link (for example Lost job or Went hungry). This means creating a new factor or reusing an existing one. Usually we create these factors inductively as we code, and revise and review and consolidate them as part of the process, as with any other kind of qualitative content analysis.

To code a causal link,

- With your mouse, highlight a piece of text within the statement which makes a causal claim. Your selection must remain within one statement and must not cross into another statement.
- Watch how that passage is copied for you into the "Quote" window. (Usually, you don't need to think about this window: you can edit the text if you really need to but it **has to remain an exact quote of one part of the text**.)
- Start to type the name of the influence factors at the**start** of the link(s) which you are going to make, in the first drop-down menu.
- If there is an existing factor which matches what you want, you can select it.
- Otherwise, you will create a new factor with the contents of what you have typed; finish what you have typed with a comma or a tab character if you want to continue to select or create another factor.
- If you want to create more than one link, you can select or create additional factors in the same box (as shown in the video below).
- When you have finished, press Enter.
- Repeat the process in the other box to specify the factors at the**end** of the link (or ends of the links).
- Press the green Save button which is now active.
- The link is created in the Map window.
- When you have finished coding one source, click the right arrow in the source navigator to code the next source.

Source Text Viewer

✨ **What you can do here:** Read your source documents and create causal links by highlighting text. When you highlight a passage that claims or implies that one thing influences or causes another, a popup lets you identify the cause and effect. This is where you do the core work of mapping out causal relationships from your source material, a process which we call *coding* The text viewr shows full text from the selected source. If you have selected multiple sources, it shows the text from the first selected source. You can:

- Highlight sections to identify causal claims. Highlighting opens the causal link editor.
- Examine and edit existing highlights by clicking on them.

Inside the header, there is an info ℹ️ icon which toggles open/shut a panel beneath it which shows the values of the custom columns for the current source e.g. gender etc.

Navigation Controls

Navigate sources:

- **Previous source**
- **Next source**

Navigate highlights within the current source:

- **First highlight in source**
- **Previous highlight**
- **Next highlight**
- **Last highlight in source** (useful if you haven't finished coding the whole text yet and want to see the last highlight)

Source selection is filtered through the sources selector dropdown. When multiple sources are loaded, the first source is displayed. The next/previous buttons cycle through sources by updating the sources selector to show the next/previous source.

This is convenient because usually when coding you will want to view the Map and Links for the same source on the right.

Clicking these buttons means that if you previously had a multiple selection, you now have only one.

Dealing with long documents in the source text viewer

For documents longer than ~30-40 pages, the text viewer automatically splits content into manageable chunks for better performance. Navigation controls appear in the "Source text" header:

- **Dropdown selector** - "Chunk 1 of 5" becomes clickable to jump to any chunk
- **Arrow buttons** - Navigate to previous/next chunk
- **"Next chunk" button** - Appears at the end of each chunk (except the last)

Visual Highlighting

Each section of coded text, each causal claim, is shown with a highlight.

For overlapping or identical highlights with multiple links, overlaps are shown with varying color opacity. Clicking on multiple highlights shows a link selector for each section.

- Multiple highlights shown with varying color opacity
- Click on overlapping highlights to select specific links

Link Editor screen

Opens when you highlight text or click on existing links.

Fields:

- **Cause and Effect selectors** - Unified project-wide label list sorted by frequency
- Both dropdowns use the same suggestions (combined causes+effects from the entire current project, not just the current source)
- Large projects: switches to type-to-search mode (min 2 chars), returning up to 200 matches; top ~300 are preloaded for quick access
- You can type new labels or select from the dropdown. You can type and select more than one cause and/or effect.
- **Quote field** - Editable text that gives the evidence for the causal claim. Also supports ellipses like this: `Actual quote [this text is ignored] quote continues blah blah.`
- **Chain toggle** - Defaults to unchecked on every fresh open/edit. If checked, saving keeps the overlay open and loads the previous Effect as the next Cause; if not checked, saving closes the overlay. The toggle remains checked only when the overlay stays open due to chaining.
- **Plain coding toggle** - Used when you want to record something which is not explicitly mentioned as a cause or an effect. Defaults to off. When on,
- the tag `#plain_coding` is added (if not already present) to the comma-separated list of tags.
- Whenever the tag `#plain_coding` is present:
  - the toggle is switched to on
  - the effect factor selector is forced to have the same contents as the cause factor
  - the effect factor selector is disabled for the user
- **Tags field** - Add tags to the link like `#hypothetical` or `check`
- **Favorite buttons** - Heart, exclamation, star toggles for marking important links or useful quotes. Later you can use these tags and favorites in filters.

Actions:

- **Save** - Create the link(s)
- **Delete** - Remove existing link
- **Cancel** - Close without saving

Links in Causal Map only have one cause and one effect. You can add multiple causes and/or effects to the boxes, and the system createsall combinations when saving. So if you put `unemployment` and `violence` in the Cause box, and `stress` and `worry` in the Effect box, the system will create four links.

About the factor label dropdown menus

By creating links, you also create the names of your factors.

In Causal Map, a factor*is*its label. Once you create a label, there is nothing else to add.

Factor names which contain semicolons ; get special treatment as they separate the different parts of 🚀 Hierarchical factors .

After beginning to create links between factors, already-coded factors will appear in the dropdown menus in the to and from factor boxes. For added convenience. The most frequently coded factors will appear at the top of this list

#doubtful? #hypothetical? Adding link tags

Link tags

Link tags are available as a special kind of memo when coding a link: you can use them to provide any kind of additional information.

There is no need to actually use a hash # at the start of a link tag, though you can if you want. Just use any unique single word which is easy to search and filter on, like #nutrition or nutrition# or nutrition–.

As usual in Causal Map, you can apply one or more tags, and you can either select existing tags or create new ones on the fly.

Later, you can filter the map (see ✨ Transforms Filters: Include or exclude tags) to show only links containing or beginning or ending with specific hashtags (or parts of hashtags), and also for links which do*not*contain specific hashtags or parts of hashtags.

You can also use tags to narrow down your searches in 🔗 The Manage Links tab.

You can display tags on your map.

Conceptually, there are two kinds of tag.

Ordinary link tags

You can use any tag which does not begin with a ? to record any other information about the link, e.g.:

- respondent doesn't like this connection
- respondent feels good about the outcome
- for you, the analyst, e.g.
  - respondent is answering a different question
  - to tag links you want to come back and review.

Weak tags

Weak tags are a special kind of tag. They are*caveats*. If you use weak tags, you should make sure that by default your maps do not include any link with a weak tag.

This is just a convention, it makes no difference to the Causal Map app.

They begin with `?` and are used to mark any link which you are not sure is always valid across the global context for the whole global map, for example:

- **the causal connection** is only valid for a specific context, e.g.
  - the respondent says this is only true for their village, not for other villages e.g. `?village X`
  - a link is only projected for the future e.g. `?future`
- you are unsure about **the claim about the causal connection**
  - a link is only a hypothesis e.g. `?hypothetical`
  - you as the analyst are not confident in the claim e.g. `?doubtful`
  - the source themselves are not sure e.g. `?source seems unsure`
  - to add other qualifying information e.g. `?probably hearsay`
  - to mark the fact that a connection is **weak or non-existent**, e.g.
    - Respondent makes a substantive claim that X does *not* influence Y, e.g. `?zero influence`
    - Respondent makes a substantive claim that X only insignificantly influences Y, e.g. `?weak`

AI Coding

Requires AI subscription

- **Model dropdown** - Select AI model
- **Prompt box** - Enter coding instructions
- **Add source prompt** - Toggle, default ON
- **Response displays** - View AI responses and full JSON

Motivation for source prompt: it is just to describe the context/background info about each source. Not necessary e.g. where all the sources are from the same context which can be described in the main prompt. But important where some differ, e.g. mid-term reports or whatever.

If Add Source Prompt is ON, then show a text area above #text-viewer-content with usual greenish Save button to edit the corresponding source prompt for the current source.

Additional controls hidden behind gear icon (experimental):

- **Temperature slider** - Control randomness (default 0)

Iterative Processing: If your prompt contains lines with `==<mark>` on their own, each section before and after the line is treated as a separate iteration. Line endings and surrounding spaces are tolerated (CRLF/whitespace OK). First iteration is normal; subsequent ones include the full prior conversation history (all previous User prompts and AI responses) to build on earlier results. Only the results of the last iteration are added to the links table; all iterations are logged in the responses panel.

Workflow:

- Select one (or more) sources to process using the sources dropdown
- Select "Skip coded sources" if you don't want to recode sources which have already been coded
- Toggle "Add source prompt" to append the new Source Prompt field before the beginning of the main prompt
- Click "Process Sources" button
- Confirmation dialog shows model, word count, and warnings
- Pre-modal quick estimates are time-boxed (~2s). If the fetch is slow, the modal still appears and estimates may show as n/a. Heavy work only starts after you click Proceed. See `webapp/js/ai-manager.js` near `AI_DEBUG_QUICKSTATS_START`.
- AI processes sources in batches
- Results are also logged to the responses table on the right of the screen

Tips on using the prompt history

- Timeouts: per-iteration budget scales by model and iteration count (cap 540s total): Flash 120s/iteration; Pro 270s/iteration.
- Concurrency: Radio group labeled "Concurrency" (1–5) next to Region in AI settings. Default 1. Increase if you want faster processing but may risk 429/timeouts.
- Logging & Responses: each chunk inserts a pending row in `ai_logs` (status pending) and updates to success/error on completion; Responses tab auto-refreshes as rows update.

Analysis Filters Filter links tab

Do qualitative causal analyseis on the selected links by filtering or manipulating them.

🔍 **What you can do here:** Apply filters to focus your analysis on specific aspects of your data. You can trace causal pathways, group similar concepts, filter by themes or demographics, and much more. Think of this as your analysis toolkit - combine different filters to explore your data from different angles.

The Filter System: overview

Use filters to narrow down and/or transforme the links you want to study. Filters are applied in order, from top to bottom.

- **Default filter**: Factor Label Filter
- **Add Filter** lets you insert filters at the start or between existing ones
- **Enable/Disable** toggles turn individual filters on or off
- **Remove** deletes a filter
- **Collapse** hides a filter's controls to save space
- **Clear All** resets to the Factor Label Filter

Hard vs Soft recoding

Most filters leave factor labels untouched, but these 'Transform filters' filters temporarily relabel factors:

- Zoom
- Collapse
- Remove Brackets
- Soft Recode Plus
- Auto Recode
- Soft Relabel
- Cluster

No filters actually change your original coding.

💡 Tip: If you want to permanently rename or "hard recode" your factors, there are several ways to do that:

- Search and replace factors
- Search and replace links

For example, after clustering (which may give labels like C11), click a factor on the map and rename it (e.g., "Wellbeing") to save the new name permanently.

Zoom Filter

🔍 **What this does:** Simplify complex factor labels by zooming to higher levels of a hierarchy. For example, turn "Health; Mental Health; Depression" into just "Health" (level 1) or "Health;

Mental Health" (level 2). Perfect for getting a big-picture view of your data.

- **Radio buttons** for levels (None, 1-9). Combine with Collapse Filter for label cleanup.
- **Level 1**:
- "foo; bar; baz" becomes "foo"
- "foo; bar; baz" becomes "foo"
- **Level 2**:
- "foo; bar; baz" stays the same
- "foo; bar; baz" becomes "foo; bar"
- **None**: No transformation

Collapse Filter

🏷️ **What this does:** Merge similar factors under one common label. Type or select multiple similar terms like "money", "income", "salary" and they'll all be replaced with the first term. Great for cleaning up data where the same concept is described in different ways.
Widgets:

- **Selectize dropdown** with existing labels where you can select one or more existing factor labels, or type parts of existing labels.
- **Matching options**: Start / Anywhere / Exact
- **Separate** toggle for individual replacements. When off, this filter replaces all matches with first search term. When on, a separate factor is created for each of the search terms.

Remove Brackets Filter

🧹 **What this does:** Clean up your factor labels by removing text in brackets. For example, "Education (primary school)" becomes just "Education". Choose between removing content in round brackets ( ) or square brackets [ ].

- **Radio buttons**: Off / Round / Square brackets
- Removes all text within selected bracket type

If you want to remove both kinds of labels, simply create another `Replace brackets` filter beneath this one.

Factor Label Filter

🎯 **What this does:** Show links connected to factors you care about (e.g. "Education"). Choose how many steps to look upstream (causes) and downstream (effects).
Widgets:

- **Factor selector** with existing labels. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from the entire project instead.
- **Steps Up** (0-5): How many levels upstream to include
- **Steps Down** (0-5): How many levels downstream to include

- **Source tracing toggle**: Retain only links which are part of complete paths which all belong to the same source
- **Highlight toggle** (default: on): Show/hide custom highlighting (⭐ star and magenta border) for matching factors
- **Matching**: Start / Anywhere / Exact. Matching is case-insensitive.

How to use: 1) Select one or more factors.
2) Set Steps Up/Down to widen or narrow the neighbourhood.
3) (Optional) Turn on Source tracing to require paths from a single source.
4) (Optional) Turn off Highlight to hide the custom highlighting. 5) The map and tables update to show only links on those paths.

All the label and tag filters including exclude filters have three radio buttons below the selectize input called Match: Start (default), Anywhere or Exact to control how search terms match against labels/tags:

- **Start**: Match only at the beginning of text (default)
- **Anywhere**: Match anywhere within the text
- **Exact**: Match the entire text exactly

Multiple search terms are treated as OR not AND. preserve and highlight factors matching ANY of the search terms.

Focused factors show with colored borders in the map and have a star added for easy identification (when Highlight toggle is on).

Exclude Factor Label filter

🚫 **What this does:** Remove unwanted factors from your analysis. Type factors like "Unclear" or "Other" to hide them from your map and tables. Useful for cleaning up your data by removing vague or irrelevant categories.

- **Factor selector** for factors to exclude. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from the entire project instead.
- **Matching options**: Start / Anywhere / Exact
- Multiple entries combined with AND logic
- If you want to exclude both/all of two or more entries, add another Exclude Factor Label filter.

Path Tracing Filter

🏞️ **What this does:** Find causal pathways between two specific points. Set a starting factor (like "Poverty") and an ending factor (like "Poor Health") to see all the causal chains that connect them. Great for understanding how problems and solutions are linked.

- **From selector** for starting factors. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from

the entire project instead.

- **To selector** for ending factors (results visible in Map and Links). Uses the same label source as the From selector (controlled by the **Show All** toggle).
- **Matching options**: Start / Anywhere / Exact
- **Steps** (1-5): Maximum path length
- **Thread tracing toggle**: Require only paths within same source
- **Highlight toggle** (default: on): Show/hide custom highlighting (⭐ star/magenta border for From factors, 🎯 target/dark yellow border for To factors)
- **Only indirect links** (default: off): Remove all direct links from From to To (only makes sense when both From and To are non-empty)

Exclude self-loops Filter

You can exclude self-loops from the maps, but that is more of a visual change. This is a real filter as part of the filter pipeline. For example, if you are using a filter like Link Frequency that might be retaining link bundles which are actually self-loops, so you might get unexpected results if you use the map setting to remove the self-loops. So this filter is a better way. It simply removes all links which are self-loops from the links table.

Link Tags Filter

🏷️ **What this does:** Filter your analysis by the tags you've added to links. Show only links tagged as "#important" or "#policy" to focus on specific themes or types of relationships you've identified.

- **Tag selector** with existing link tags from current project
- **Matching options**: Start / Anywhere / Exact

Combine Opposites filter

🔄 **What this does:** Unify opposite factor labels by matching tag numbers. If you have pairs like `Foo [99]` and `Bar [~99]` (where `~99` indicates the opposite), this filter rewrites `Bar [~99]` as `Foo [99]` to combine them under one label. The `flipped_cause` and `flipped_effect` columns track which causes and effects were flipped.
Toggle – Turn the filter on/off.

Strip tags from labels (default: on) – When enabled, removes `[N]` and `[~N]` tag patterns from labels after combining opposites. This keeps labels clean while preserving the tracking information in the `flipped_cause` and `flipped_effect` columns.

Labels can be written in pairs like:

- `Foo [99]`
- `Bar [~99]`

where `Bar` represents the opposite of `Foo`. The square brackets are optional - you can use `Foo 99` and `Bar ~99` - but brackets make it easier to remove tags later using the Replace Brackets filter.

If there are any such pairs, with matching integers, and the filter is switched on:

rewrite any Bar [~99] filters as Foo [00] and add new columns:

- `flipped_cause` column tracks which causes were flipped
- `flipped_effect` column tracks which effects were flipped

to the current augmented links table, so that if the label has been flipped, the value is True and otherwise False.

Wire up the filter as part of the standard filter system with save/restore to URL etc.

Also, when calculating new links table, create new text columns:

- `source_count_with_opposites`
- `citation_count_with_opposites`

The embellished counts always show all variants with custom SVG icons (no total prefix). Four circle icons represent the flipped status:

- ‾ (unflipped/unflipped)
- ╲ (unflipped/flipped)
- ╱ (flipped/unflipped)
- _ (flipped/flipped)

So if a bundle has 12 citations where 5 are unflipped/unflipped, 2 have flipped cause and flipped effect, and 1 has flipped cause but non-flipped effect, the text is: `light-blue-circle5, dark-red-circle2, mixed-circle1`. If nothing were flipped, the label would just be `light-blue-circle12`. Do the same with source counts too, counting the unique sources in each variant.

When the filter is on, and source count or citation count is selected, the graphviz and graphviz maps change to use `source_count_with_opposites` or `citation_count_with_opposites` just for the labels. The edge width calculation remains driven by source count or citation count, as selected.

Exclude Link Tag filter

🚫 **What this does:** Remove specific types of links from your analysis. Exclude links tagged as "#uncertain" or "#duplicate" to focus on higher-quality data. Helpful for filtering out questionable or irrelevant causal claims.

- Same as Link Tag filter except *exclude* links containing these tags. Multiple entries are combined with AND, i.e. only exclude links where both entries match. (💡 Tip: if you want to exclude both/all of two or more entries, add another filter).

Exclude self-loops Filter

You can exclude self-loops from the maps, but that is more of a visual change. This is a real filter as part of the filter pipeline. For example, if you are using a filter like Link Frequency that might be retaining link bundles which are actually self-loops, so you might get unexpected results if you

use the map setting to remove the self-loops. So this filter is a better way. It simply removes all links which are self-loops from the links table.

Link Frequency Filter

📊 **What this does:** Focus on the most important causal relationships by filtering out rare ones. Choose "Top 10" to see only the most frequently mentioned connections, or set a minimum threshold like "at least 3 sources" to ensure reliability.

- **Slider** (1-100) for threshold
- **Type**: Top vs Minimum
- **Count by**: Sources vs Citations

Examples:

- **Minimum 6 Sources**: Only links mentioned by 6+ sources
- **Top 6**: Only the 6 most frequent link bundles

By default, setting the slider to 6 means we are selecting only links with at least 6 citations.

If you switch to “Sources”, we are selecting only links with at least 6 sources.

If you switch to “Top” we are selecting only the top 6 links by citation count, etc. The selection respects ties, so that if there are several links with the same count, either all of them or none of them will be selected.

Factor Frequency Filter

📈 **What this does:** Similar to Link Frequency, but focuses on the most important factors (causes and effects). Show only the most frequently mentioned themes or concepts to identify the key issues in your data.
Same controls as Link Frequency but applies to factors instead of links.

Source Groups filter

👥 **What this does:** Filter your analysis by participant demographics or document characteristics. For example, show only responses from "women aged 25-35" or interviews from "urban areas." Perfect for comparing how different groups see causal relationships. This is very similar to the [Source Groups widget](../source-groups-sub-panel/) in the Sources sub-panel, but having it here too means you can add multiple source filters to the pipeline.

- provides
- a prepopulated dropdown called Field with all the metadata fields plus title and projectname
- another multi-selectzie called Value. Multiple values work as OR: either/any count as a match
- a previous/next button pair to cycle through values of the selected group
- Example: Add two Source Groups filters in the pipeline to combine criteria (e.g., first filter Field = gender → Value = women, then another filter Field = region → Value = X) so you see links from women AND from region X.

Everything Filter

👥 **What this does:** Filter your analysis by any characteristic of your links and their sources. Useful for anything not covered by the other filters, for example, - show only links with negative sentiment - Show only links from one source Also shows source separators and their values, often used for common sections within multiple sources texts.

- **Field dropdown** with all fields in the links table
- **Value selector** filtered by selected field
- **Navigation buttons** to cycle through values
- **Clear button** to reset

Soft Relabel Filter

👥 **What this does:** Temporarily relabel factors.

- **Old factor labels** listed on the left
- **New factor labels** editable, listed on the right
- **Load labels button** when pressed, adds into the Old labels list any current factor labels (in links as currently filtered) which are not yet listed in the Old labels list and adds the same Old label to the New column as default.
- **Clear button** to clear the New fields
- **Clear ALL button** to clear all rows

Effect: all factors exactly matching any of the labels in the Old list are relabelled with the corresponding labels from the New list. factors not listed are not relabelled but preserved.

Many use cases:

- temporarily merge multiple factors into one
- you are using magnets and you can't really use the formulation you want because you want to maximise similarity with existing labels
- eg you are using "floods" as a magnet but you really want it as a hierarchical factor like "environmental problems; floods" but you can t use that as a magnet.

Keyboard shortcuts (Win/Linux ⇄ macOS):

- Tab / Shift+Tab: move focus down/up between NEW cells
- Arrow Up/Down: move focus up/down between NEW cells
- Alt+Arrow Up/Down (mac: Option+Arrow): move the current row up/down
- Ctrl+Arrow Up/Down (mac: Cmd+Arrow): move the current row up/down
- Delete current row:
- Shift+Delete (mac: Shift+Fn+Backspace) or
- Ctrl+Shift+K (mac: Cmd+Shift+K)

Potentially, one NEW label might have multiple OLD labels.

Soft Recode Plus filter

Requires AI subscription

🧲 **What this does:** Group messy factor labels under clearer names you choose (called "magnets"). Example magnets: `Improved health`, `Education programs`, `Income changes`. The filter finds the closest magnet for each label and replaces it.

Controls:

Create Suggestions for Magnets

(collapsed by default): Optional. Ask AI to propose clear names from your current labels. Insert adds them to your magnets box to review/edit.

- **Number of clusters** – Choose how many groups to find for AI suggestions.
- **Labelling prompt** - With the usual buttons to save and recall previous prompts
- **Insert**

Main panel

- NEW: **Only unmatched** – A new toggle which appears right at the top, before the Create Suggestions subpanel. default off.
- **Magnets** – One magnet per line. Saved per project. Use Prev/Next to browse recent sets.
- **Similarity slider** – The raw labels are dropped if they are not at least this similar to at least one cluster.
- **Drop unmatched** – If on, remove links whose labels don't match any magnet. If off, keep them as they are.
- **Save** – Save magnets and apply the recode.
- **Remove hierarchy** – Strip any text before the final semicolon
- **Clear / Prev / Next** – Manage saved sets.
- **Recycle weakest magnets:** – A slider starting at 0 , default is 0. If the slider is n >0, then we look at the cluster assignenments which would have been returned and find the n clusters which we are going to recycle. Reassign them to their nearest cluster, providing the similarity is still better than the similarity cutoff. This way we don't lose factors / links which are otherwise assigned to smaller clusters which may get excluded later on in the filter pipeline. When it is on zero, it makes no difference and we just use the solution based only on the magnets, similarity, and remove_hierarchy. The maximum value changes to match the total number of magnets.

Recoded columns

When you use Soft Recode Plus, the Links and Factors tables show special columns that track which labels have been recoded:

- **Links table**: Shows `_recoded_cause` and `_recoded_effect` columns (✓ for recoded, ✗ for not recoded)
- **Factors table**: Shows `_recoded` column (✓ if the factor appears at least once as recoded, ✗ otherwise)

- These columns only appear when Soft Recode Plus is active in your filter pipeline
- You can filter by these columns using the True/False dropdowns in the table headers

These columns track recoding from any filter that transforms labels: Soft Recode Plus, Zoom, Collapse, Remove Brackets, Soft Relabel, Cluster, Hierarchical Cluster, and Combine Opposites.

Process only unmatched NEW

the point of this is: what if I apply some (maybe standard) magnetisation and matches plenty of factors but there might be some important material left unmatched, not just noise. so i can use a PAIR of these filters. in the first one, I leave OFF its Discard Unmatched toggle and in the second filter switch ON its Only Unmatched filter. (if there is no preceding SRP filter with Discard Unmatched=OFF, this second SRP filter does nothing).

So now,

- the Create Suggestions (if used) optionally processes ONLY the UNMATCHED factor labels
- the magnetisation (if labels are non-empty) works only on the unmatched factor labels.
- the actual output of the second filter is now the union of both soft-recode processes, i.e. the original matches from the first and the new matches of the previously discarded material from the second.
- the Discard Unmatched on this second filter works as usual: if it is OFF, then we also return all the still-unmatched labels.

Meaning Space (2-D embeddings)

Go to the map formatting and select Layout → Meaning Space to see a 2-D scatter of your factors in “meaning space”.

- Magnets are shown with labels; raw factor labels are dots.
- Colour indicates the magnet group; magnet dot size represents group size.
- You can pan (drag) and zoom (mouse wheel and zoom controls).
- Double-click on an empty part of the map to zoom in at that point.
- Tooltips on dots show the original (raw) labels and the magnet label.

Motivation for Remove Hierarchy

"Remove hierarchy", default off. if on, strip any text before a final semi-colon, if no semi-colon, do not change the text.

```
something; another thing
```

is treated same as

```
another thing
```

.... but it continues to be treated as "something; another thing" in the rest of the filter pipeline.

Quick workflow:

1) (Optional) Open Create Suggestions for Magnets panel → set Number of clusters and use Insert to get AI suggestions.
2) Use these suggestions and/or edit them, paste or type your own magnets (one per line).
3) Click Save.

- Clusters your current labels (factors as currently filtered), ranks typical examples, and asks AI to suggest clear names.
- Returns suggested names into the magnets box; you can edit them before Save.

See tips on using the history to reuse both your labelling prompt and magnet sets.

Motivation for "recycle weakest magnets": suppose you create 20 magnets, and then apply more filters like say a link frequency filter so that you end up with say only 5 factors. If you then *remove* those factors from the magnets list which are *not* included in the final output, you will usually increase the coverage of your map (re-assigning raw labels which fit best with one of the "lost" labels but still fit well with one of the "surviving" labels). This is what the Recycle slider does: it recycles the specified number of smaller magnets and reassigns them to the larger magnets. So in the example, if you start off with 20 magnets but your final map only shows 5, try recycling say 10 or even 15 of the missing factors.

Note that Recycle Weakest Magnets is applied BEFORE Drop Unmatched.

Clustering filter

Requires AI subscription

🎯 **What this does:** Automatically discover groups of similar factors in your data using machine learning. The system finds natural clusters of related concepts and labels them with cluster numbers. Great for exploratory analysis when you're not sure what causal themes exist.

- **Enable toggle** (starts disabled)
- **Number of clusters** (1-9)
- **Server-side processing** using `cluster_factors_pgvector` database function
- Uses k-means clustering on factor embeddings
- Labels clusters with numeric IDs

Auto Recode filter

🪜 **What this does:** Quickly turns your current set of labels (after any previous filters like Zoom) into a simple tree you can "roll up" or "open out". Pick a small number of clusters for a big-picture view, then nudge the Balance and Similarity to tidy results. Designed for fast, practical exploration on real projects.

Motivation

Making sense of hundreds or thousands of factor labels is hard.

You might use something like soft Recode Plus, but often you'll ask for 20 clusters to cover a wide range of meanings. Then after filtering out insignificant data, you end up with only 7 clusters — losing coverage. Ideally you'd go back and recreate just 7 clusters, but that gives different results. Frustrating!

The point of this Auto Recode filter: have your cake and eat it. Ask for an foldable/unfoldable hierarchical solution. When you move the slider to 15, you get the best solution for 15 clusters. Slide it to 3, you instantly get the best solution for 3 clusters.

Controls:

- **Enable toggle** (starts disabled)
- **Balance (0..1)**: 0 = prefer more distinct clusters; 1 = prefer more even sizes. Changing this can be slow because the tree has to be rebuilt
- **Number of clusters (K)**: 2–50. Unfolds the returned tree locally to K. This is fast unless you increase beyond 20.
- **Similarity ≥**: prune locally by similarity to the centre of each cluster.

NEW: **AI labelling prompt** with history controls. Use this to suggest clearer names for each cluster:

- Saved in the prompts table as type `hierarchical_label` (shared across projects; history shows current first then others).
- A Save button stores your prompt; it also auto-saves on blur and after the first tree build.
- When you raise K (unfold deeper), we call AI in parallel only for the two new child clusters introduced by each applied split, using up to 8 representative labels per child as context. For K clusters this is K−1 requests. Folding to fewer clusters does not call AI; existing AI labels or medoid representatives are used.
- If the prompt is blank, we show the medoid representatives for each cluster.
- If earlier splits already have AI labels (K > 1), we include a reference list of those labels so new labels avoid overlapping meanings.

NEW: **Seed labels (optional)** with history and strength:

- Provide up to K seed labels (one per line). Seeds softly influence split formation but are not included in the final tree (not nodes, not representatives).
- Saved in the prompts table as type `hierarchical_seeds` with standard history controls (Prev/Next/Dropdown/Save).
- Seed strength (0..1) adjusts influence; 0 is a no-op (identical to no seeds). Changing strength or seeds triggers a single backend rebuild (like Balance). Changing K or Similarity does not re-call the backend.

How to use (quick):

- Add the filter and enable it. We build a quick draft tree from the labels you see now (respecting any filters above, like Zoom).

- Set **Balance** if you want more equal-sized groups; the first build may take a moment on large projects (one server call).
- Use **K** to choose how many clusters to show. Changing K is instant (no extra server calls).
- Use **Similarity ≥** to drop weak matches. If either side of a link isn't matched, that link is hidden.

Notes:

- On very large projects, we automatically sample a representative set to build the tree, then assign the rest to the nearest cluster. This keeps things responsive while preserving the overall picture.
- 💡 Tip: changing the number of factors should be instant if they are less than 20. Setting more than 20 can be slow. If you are going to want more than 20, set this number initially to the maximum number you are likely to want. You can then easily reduce it. Gradually decreasing the number is fine, but *gradually increasing* it will be very slow.

A good prompt looks something like this:

> This is a list of many raw labels grouped into two different clusters, with their cluster IDs, together with a reference list of other labels. Return a list of two new labels, one for each cluster ID. Each label should capture the meaning of the whole cluster, using similar language to the original raw labels, but in such a way that the labels you create are distinct from one another in meaning. Try not to be too generic, try to be as concrete as you can. Do NOT provide labels which include causal ideas, like "X through Y" or "X leading to Y" or "X results in Y" or "X improves Y" etc. Equally, don't include conjunctions in the title like "X and Y". The meaning of the labels you give me should ideally not overlap in meaning with one another or with the labels in the reference list.

Optimized Cluster filter

⚠️ DEPRECATED **Requires AI subscription**

> ⚠️ **This filter is deprecated.** Its functionality has been merged into Soft Recode Plus. The filter still works for backward compatibility with existing bookmarks/URLs, but new instances cannot be created. Use Soft Recode Plus instead for optimal clustering and recoding.

🎯 **What this does:** Automatically finds the most optimal factor labels to use as centroids through genuine optimization. Unlike regular clustering that just groups similar items, this finds the best possible n≤N labels that maximize coverage with similarity ≥S. Perfect for discovering the most representative concepts in your data.
Controls:

- **Max Centroids (n)** - Maximum number of optimal centroids to find (2-50)
- **Similarity ≥** - Minimum similarity threshold for grouping labels (0-1)
- **Timeout (s)** - Optimization time limit in seconds (5-60)
- **Drop unmatched** - Remove labels that don't meet similarity threshold
- **Real-time status** - Shows optimization progress and results

How it works:

1. Extracts all unique labels from your current data (1K-30K labels supported)
2. Runs iterative optimization with multiple strategies (random, frequency-based, diverse selection)
3. Uses hill-climbing optimization to find the best possible centroids
4. Shows coverage percentage and timing information
5. Returns recoded links table with optimal centroid labels

Optimization Strategies:

- **Random selection** - Tests random starting points
- **Frequency-based** - Prioritizes most connected labels
- **Diverse selection** - Maximizes distance between centroids
- **Hybrid approach** - Combines best-so-far with random exploration

Performance Features:

- **Sampling strategy** for datasets >1000 labels (uses representative subset)
- **Early termination** when excellent coverage (≥95%) is achieved
- **Configurable timeout** prevents infinite optimization loops
- **Multiple iterations** with different starting strategies for robustness
- **Smart caching** - Embeddings cached separately from algorithm parameters for fast parameter changes
- **Quote-safe processing** - Handles labels with quotes, apostrophes, and special characters

Technical Implementation:

- Client-side optimization using cosine similarity on embeddings
- Hill-climbing algorithm with local search improvements
- Genuine optimization problem solving (not just k-means clustering)
- Real-time UI feedback showing progress and final results
- Handles massive datasets efficiently through smart sampling
- **Original label preservation** - Stores original labels in `_recoded` metadata for map display
- **Chain compatibility** - Works seamlessly with zoom filter and other transformations

Soft Recode Integration:

- Optimized cluster results available as magnet source in Soft Recode filter

- AI can generate meaningful labels for optimal centroids
- Seamless workflow from optimization to AI-powered naming

This filter implements the optimization challenge described in the technical documentation: finding optimal centroids that maximize label coverage within similarity constraints.

Tribes filter

Requires AI subscription

🏛️ **What this does:** Group your sources (participants/documents) by how similarly they describe causal relationships. This reveals different "tribes" or perspectives in your data - for example, optimists vs. pessimists, or urban vs. rural viewpoints.
Controls:

- **Number of clusters** - Radio buttons: Off, 1-9
- **Similarity cutoff** - Slider: 0-1
- **Drop unmatched** - Toggle
- **Min cluster %** - Slider: 0-20% (prevents "1 big + many singletons" pattern)
- **Counts (Report)** - For the Tribe Report tables: count by **Sources** (unique participants/documents) or **Citations** (links). Default: Sources.

.

It returns:

- `tribeId` (cluster ID)
- similarity to the centroid
- similarity rank These are joined to the links table by source ID and appear as additional columns. If Drop unmatched is ON, links with similarity below the cutoff are removed.

We can then show maps for each tribe and/or for the most typical source in each tribe. we could also then create a typical story centred around the current factors, i.e. told in terms of our concepts.

--->

Custom Links Label

🏷️ **What this does:** Configure how link labels appear on your map based on any field in your data. Choose what information to display (like tribe memberships or custom attributes) and how to show it (counts, percentages, or statistical significance).
Controls:

- **Field** - Dropdown of available fields from your filtered data (typically shows custom fields like tribe ID)
- **Counts** - Choose whether to count **Sources** (unique participants/documents) or **Citations** (links)

- **Display mode** - Choose how to show the data:
- **Tally** - Show counts for each value (e.g., "T1:4 T2:3")
- **Percentage** - Show what % of each value's total links appear in this bundle (e.g., "T1:34% T2:22%")
- **Chi-square** - Show bundle size, then which values are significantly over-represented (⬆) or under-represented (⬇) (e.g., "45 (T1⬆ T3⬇)")
- **Chi-square (with counts)** - Also show the observed count for each significant value (e.g., "45 (T1 4⬆, T3 3⬇)")
- **Chi-square (with counts/totals)** - Also show observed/total for each significant value (e.g., "45 (T1 4/5⬆, T3 3/6⬇)")

To use:

1. Add the Custom Links Label filter to your pipeline
2. Select a field (e.g., `custom_tribeId` after running the Tribes filter)
3. Choose a display mode
4. In Map Formatting, set Link Labels to "Custom Links label"

Example use cases:

- **After Tribes filter:** Show which tribes contribute to each connection (T1:5 T2:2 T3:1)
- **Significance testing:** Identify connections where certain tribes are surprisingly over/under-represented (T1↑ T3↓)
- **Custom attributes:** Display any custom field you've added to your data

<!---

Technical details:

This is a **non-filtering** filter - it doesn't change which links appear, only configures how they're labeled on the map.

Display modes: All modes use the **Counts** toggle:

- **Citations**: each link counts as 1 observation
- **Sources**: each unique `source_id` counts as 1 observation (per value and per bundle)

- **Tally:** For each value, show its count within the bundle (based on chosen Counts unit)

- **Percentage:** For each value, calculates: (count in this bundle) / (total count of this value across all filtered links) × 100 (based on chosen Counts unit)
- **Chi-square (no counts):** For each value, tests whether observed differs from expected (based on chosen Counts unit):
- Expected = (bundle size) × (value total) / (grand total)
- Chi-square contribution = (observed - expected)² / expected
- Critical value for p < 0.05 with df=1 is 3.84

- Format: `bundleSize (value↑, value↓)` (only significant values shown)
- **Chi-square (with counts):** Same test, but format includes observed count: `bundleSize (value observed↑, value observed↓)`
- **Chi-square (with counts/totals):** Same test, but format includes observed/total: `bundleSize (value observed/total↑, value observed/total↓)`

The filter populates its field dropdown from `currentFilteredLinks` (the output of the filter pipeline), so it sees all fields added by previous filters.

Projects Panel

📁 **What you can do here:** Organize and manage all your research projects in one place. Create new projects, share them with collaborators, add descriptive tags, and control who can view or edit your work. You can also merge multiple projects together or archive old ones to keep your workspace tidy. See also the [File menu](../file-menu/) for more project management options.

Project Management

- **New project** - Create with name and description
- **Load project** - Open selected project
- **Edit project** - Click row to modify name, description, tags, sharing (or use Edit project Modal)
- **Archive/Unarchive** - Hide/show projects (archived projects are read-only and hidden from public view)
- **Read-only toggle** - Restrict editing even for owners/editors (independent of archive status)
- **Archive toggle** - Show/hide archived projects in the table

Bulk Operations

Select projects with checkboxes, then:

- **Delete** - Remove projects and all data
- **Apply Tags** - Add tags to selected projects
- **Remove Tags** - Remove tags
- **Toggle Archive** - Archive/unarchive
- **Merge** - Combine multiple projects into one

Sharing and Permissions

- **Email-based collaboration** : add and remove colleagues' email addresses
- **Locked / Read-only permissions** for viewing without editing
- **Global sharing** for public read-only access
- **Permission badges** next to project names
- **Admin only: admin panel** for user management

Note: In the Edit Project modal, the informational notice "Your projects are public and can be viewed by anyone. Upgrade your subscription to keep them private." is shown only to users who can edit the project (owner, editor, or admin). Viewers do not see this notice.

Versioning

The app automatically backs up your project, so you can restore earlier snapshots if you want.

- An automatic backup is made every 10 minutes if you have made changes.
- You can make a manual backup from the Project Info screen.
- You can use the Version dropdown menu in the Project Info screen to see all available backups with details.

- From here you can restore a backup, with a confirmation step before applying changes.
- After restoring an earlier version, you can always go back to the latest version if you want, using the same dropdown menu.

This panel shows a dropdown list of times when you made changes to the mapfile in UTC/GMT. Along with the size of your file which can help you identify which timepoint you want to revert to. It can be easy to forget what time you made alterations to your file, so if you're likely to want to restore a previous map it is best to note the time so that you can easily return to it.

Sources Panel

📚 **What you can do here:** Upload your research documents (PDFs, Word docs) and organize them with custom metadata like participant demographics or interview dates by adding and editing **custom columns**. This is your document library and metadata manager.

Upload & Setup

Upload Source Texts

- **Click to select** one or more PDFs or DOCX or RTF files
- **Optionally split large documents into multiple sources** using separator patterns
- **Confirmation dialog** shows projectname→ID mapping
- On completion, the app automatically:
  - selects the sources via Sources Dropdown
  - switches to the Create Links sub-tab and loads the first source's text into it
  - on the right, switches to Sources panel with and the View & Edit subpanel

Splitting documents into multiple sources with source separators

This feature helps you split individual documents you upload into multiple sources. These separators are *hard*: you use them just once, on uploading one or more documents, to produce multiple sources.

- In the Confirm Upload screen, there is a Sources Separator text box
- Text lines matching special "regex" patterns separate into multiple sources. So if you have sections marked with "Source Number 12", "Source Number 13" etc, just put "Source Number.*" in the box.
- This will produce multiple sources with "Source Number 12" etc as source_id
- Live preview of new source IDs with count
- User can leave blank for normal upload

Section Separators

This feature helps you split existing source texts into sections. These separators are *soft*: they don't permanently change the file and you can add or remove one or more separators on the fly.

- **Two-line expandable textArea** for one or more (regex) patterns
- **Section header detection** within imported texts
- **Special styling** for matching rows in text viewer
- **Links created in the different sections can be filtered using the Everything Filter**. So you can do things like "Show me all the causal claims (links) only in answers to Question 7".

- Into the box, you type special "regex" patterns to create sections withing source texts. So if you have sections marked with "Section Number 12", "Section Number 13" etc, just put "Section Number.*" in the box.

- This will highlight the sections and create fields like "Section Number 12" etc as section ID.

Sample Check

You want a table showing your sample: gender * region?

Use this simple customisable table to check your sample according to any custom columns you have defined.

- See also: Statistics Panel and Analysis Filters.

Custom Column Analysis

- **Aggregate by** multi-select for cross-tabulation
- **1 column**: Simple count table
- **2 columns**: Cross-tabulation table
- **Maximum 10 values** per column for analysis
- **URL state preserved** for bookmarking

View & Edit Your Sources

- See also: Tips for all tables and Custom columns.

Sources Table

- NEW column **Source Prompt** - this new column shows the first few characters of any text in this field. It can be edited as usual with the existing pencil icon/ edit button in each row.
- **Checkbox selection** for analysis pipeline
- **Row editing** with keyboard navigation
- **Custom columns** for metadata
- **Uncoded column** - Shows true/false for sources with no links; filterable to find uncoded sources
- **Fullscreen mode** available

Table Editing Features

- **Range selection** - Click and drag
- **Copy/paste** - Ctrl+C, Ctrl+V
- **Arrow key navigation**
- **Delete/Backspace** to clear cells
- **Column/row selection** - Click headers
- **Double-click editing**

Custom Columns

- **Manage Columns** opens a modal to add and remove multiple columns at once
- **Toggle visibility**
- **Double-click a cell in a custom column to edit it** or via source edit modal
- **Copy and paste** selections with ctrl-C, ctrl-V.
- **These columns are available elsewhere in the app, e.g.**

- In the Source Groups filter
- In the Everything filter
- In the Sample Check table

Map Panel

🗺️ **What you can do here:** See your causal relationships as an interactive network map. Drag nodes around, click on links to edit them, and use the controls to customize how the map looks. You can even drag one factor onto another to quickly create new links. This is where your data comes to life visually.

Map Controls

- **Jump to factor** - Type-to-search dropdown to quickly find and select factors on the map. Type to filter options, then press Enter to select all matching factors. Supports multiple selections.
- **Refresh layout** - Return the map to its original state before zoooming, moving etc.
- **Copy image to clipboard** - Get a very high-quality image copied straight into your clipboard which you can paste in a report or presentation.
- **Copy legend**
- **Zoom in/out** controls
- **Double-click** anywhere on the map background to zoom in to that point

Map Legend

Discrete text legend showing:

- projectname and included sources
- Citation coverage percentage
- Visual encoding explanations (link sizes, colors, numbers)
- Applied filters summary
- 💡 Tip: Click Copy legend to copy this text to clipboard.
- You can drag the legend box to reposition it on the map.

Map Formatting

Customisable formatting (Things you can tweak)

Layout: change how the map is laid out and how you interact with it.

- Interactive and most of the other layouts are good while you are conducting your research. They are fast and you can interact with the results -- moving factors around, clicking to edit, etc.
- Print/Graphviz layout is best for static images e.g. for reports and journal articles.
- Direction: For Interactive and Print/Graphviz layouts, choose LR (left-to-right, default), TB (top-to-bottom), or BT (bottom-to-top).

Factor Labels: (you can see the same data in the Factors Panel)

- Source count (default)
- Citation count
- Sentiment (mean incoming edge sentiment, -1..+1)

- None

Link Label Font Sizes Link Widths: Citation count (default) Source count, None **Link Colour:** Default link line colour (applies to Interactive and Print/Graphviz layouts). When sentiment is neutral (0), this colour is also used for arrowheads and node borders. **Link Labels:**

- **Source count** (default)
- **Citation count**
- **Sentiment** (mean edge sentiment, -1..+1)
- **Custom Links label** - Use configuration from Custom Links Label filter
- **Unique Sources** - Alphabetical list
- **All Sources** - Complete list with repeats
- **Unique Tags** - Alphabetical list
- **All Tags** - Complete list with repeats
- **None** - Show no labels on links

Factor Colors:

- **Outcomeness** (default) - Based on in-degree ratio
- **Source count**
- **Citation count**
- **None**

Factor Sizes:

- **Citation count** (default) - Font size scales with citation count
- **Source count** - Font size scales with source count
- **None** - All factors use uniform size (increased by 50% for visibility)

Self-loops:

- **Show toggle** (default: on)

Fixed visual appearance (things you can't tweak)

Link Styling:

- Arrowheads colored by mean sentiment (neutral uses Link Colour)
- Color scale: muted blue (+1) → grey (0) → muted red (-1)
- Bezier curved edges with bundling

Factor Styling:

- Size scaled by node degree (with bounds)
- Border color reflects mean incoming edge sentiment
- Matched factors show dashed colored borders

Interactive Features

These work for all layouts except Print/Graphviz layout (which is mostly for static export, but does support clicking nodes/links now).

- **Drag factors** to temporarily reposition them
- **Drag factor to factor** to create new links
- **Shift+drag** for box selection of multiple factors (opens edit modal)
- **Ctrl+drag** for box selection of multiple factors (direct selection, no modal)
- **Click a link** to edit.
- **Click a factor** to edit; shift-click or ctrl-click to add to selection without opening modal.

Editing and deleting (multiple) factors

- Select factor(s) by clicking a factor, shift-click or ctrl-click to add more, or shift+drag/ctrl+drag a box around multiple factors, then:
- Move selected factors together
- Delete matching factors everywhere or in current view only
- Rename matching factors everywhere or in current view only

What does "everywhere or in current view only" mean?

- **everywhere**: all links containing factors with exactly the selected labels will be deleted
- **in current view only**: all links containing factors with exactly the selected labels (and matching the current filters, i.e. those you can see in the current map) will be deleted

💡 Tip: By control-clicking or shift-clicking multiple factors you can easily rename several at once, e.g. you can merge multiple factors as a single factor.

Grid layout

Factors containing a tag of the form `(N.M)` or `(N,M)` anywhere in the label (where N and M are integers) are positioned on a grid layout. The grid coordinate tags are automatically stripped from displayed labels.

Grid layout toggle: Enable/disable grid layout in Map Formatting. Defaults to enabled. Disabled automatically when no grid tags are present.

Interactive Layout:

- Grid-tagged factors are positioned at their grid coordinates and locked in place
- Other factors with no grid tag are positioned freely within the grid bounds
- Grid bounds: from smallest x -1 to largest x +1, and smallest y -1 to largest y +1

Print/Graphviz Layout:

- Grid-tagged factors anchor the initial and final ranks:
- Factors with minimum rank coordinate (first number) are anchored at `rank=min` (initial rank)

- Factors with maximum rank coordinate are anchored at `rank=max` (final rank)
- This improves layout stability while allowing Graphviz to position other nodes optimally
- Grid coordinate tags are stripped from labels in the output

Grid coordinates respect layout direction:

- **First number (N)** always maps to the rank direction (main flow direction)
- **Second number (M)** always maps to the perpendicular direction
- **BT (Bottom-Top)**: First number = y (rank), y starts at bottom (flip y), second = x
- **TB (Top-Bottom)**: First number = y (rank), y starts at top (normal), second = x
- **LR (Left-Right)**: First number = x (rank), x starts at left (normal), second = y, y starts at top
- **RL (Right-Left)**: First number = x (rank), x starts at right (flip x), second = y, y starts at top

Vignettes

📝 **What you can do here:** Generate AI-powered narrative summaries of your causal maps. Choose between a "whole map" summary that covers all the relationships, or a "typical source" story that focuses on one representative case. Perfect for creating reports or explaining your findings in plain language.
How to use:

1. Select your **model** and **region** settings
2. Choose **Whole Map** or **Typical Source**
3. Enter or edit your **prompt** (use the navigation buttons to browse previous prompts)
4. Click **Write Vignette** to generate

Whole Map: Creates a summary of all relationships in your current map view. the app provides the following data which is appended to the prompt:

- The overall map (same as you can see) including factor frequencies and bundled causal links with average sentiment
- Up to 5 "typical sources" that tell the most common stories within the current map, with their quotes and metadata including source ID, Title and Filename.

Typical Source: Focuses on the single most representative source, showing individual links with quotes and sentiment.

Output format: Results are displayed as markdown with support for:

- Headers, bold, italic text
- Bulleted and numbered lists
- Callouts/quotes (using `>`)
- Code blocks

You can edit your prompt to change the tone, audience, or focus before generating. See the tips on using prompt history for more details.

Factors Panel

🏷️ **What you can do here:** See all the factors (causes and effects) in your data, ranked by how often they appear. Select multiple factors to rename them, merge similar ones, or delete unwanted entries. If you've added demographic data to your sources, you can also see statistical breakdowns showing which groups mention certain factors more often.
above both links and factors tables add a toggle "Use filters". If on (default) the table is filtered by the links filters. If off, we bypass this part of the pipeline and filter only by project and sources.

The Factors panel displays all unique labels from the current filter pipeline.

Table Features:

- Columns include:
- **Citation Count** – total number of citaions of this factor (as cause or effect)
- **Source Count** – number of different sources mentioning this factor
- **Citation Count: In** – number of citations of this factor as an effect of something
- **Citation Count: Out** – number of citations of this factor as a cause of something
- **Source Count: In** – number of sources mentioning this factor as an effect of something
- **Source Count: Out** – number of sources mentioning this factor as a cause of something
- Sorted by citation count (descending order)
- Click-to-select (no checkbox column)
- Server-side pagination consistent with other tables
- Actions column with edit button to open factor edit modal

Action Buttons:

- Delete: Remove selected factors
- Relabel: Rename selected factors
- Search/Replace: Find and replace text in factor names
- Merge: Combine multiple factors into one
- Buttons disabled until factors are selected
- 💡 Tip: Use Search/replace for quick, scoped relabeling.

Find out more about bulk delete and relabel of factors here.

Bulk factor labels editor

The search/replace functions in the factors and links tables are useful, but what if you have thousands of factors to look at? You might prefer this bulk editor.

Toggle the **Bulk Edit** switch to edit multiple factor labels at once. The table header remains visible for sorting and filtering, while the table body is replaced with a line-by-line editor.

Features:

- **Multi-cursor editing**: Use `Alt+Click` or `Ctrl+Alt+Up/Down` to add multiple cursors

- **Find occurrences**: Use `Ctrl+Alt+Right` to add next occurrence, `Ctrl+Alt+Left` for previous
- **Select all matches**: Use `Ctrl+Shift+L` to select all occurrences of selected text
- **Line-by-line editing**: You can only edit existing labels - you cannot add, remove, or reorder lines
- **Recoded labels**: Labels that have been recoded (shown with yellow background) are read-only and cannot be edited
- **Sort and filter**: The editor automatically updates when you sort or filter the table (any unsaved edits are discarded)

How it works:

1. Toggle **Bulk Edit** on
2. Edit factor labels directly in the editor
3. Press **Save Changes** to apply your edits
4. A confirmation dialog shows which labels will change and how many links will be affected
5. After saving, the editor refreshes to show the updated labels

NEW: add a second column to this div so that the editor takes up 9/12 of the width. in the new column, provide some live info about the selected factor: source and citation counts, and a list of sources mentionoing it.

Search/replace

Near the top is a row containing a search box. If you type something into it,

- a replace box and a Replace button also appear.
- the table is filtered to show only matching rows

The search is **case sensitive**.

You can then alter what you see in the Replace box:

- in the factor label column in the table, you see a preview of what the affected rows would look like;
- if you delete all the text so the replace is empty, the preview shows the effect of deleting the search text from each label.

Then when you are satisfied, check all the checkboxes where you want to update the labels as shown. If you want, select all rows using the checkbox at the top of the column. Note, if there are more hits than fit on this page of your table, you'll want to either treat each page separately or increase the page size with the Page Size selector.

Finally, hit the Replace button to actually update the labels as shown in the rows you selected. What actually happens is that the Cause and Effect labels in all the currently selected links are changed. As you'd expect, this search/replace only affects the factors for the currently selected

links: for example if you have only selected the first three sources, this update will not affect the links in the other sources.

.

Demographic Breakdowns

- **Breakdown selector** - Choose custom columns to analyze by demographics
- **Count type** - Source count (default) or citation count
- **Display mode** - Counts (default) or % of baseline (cell as a percent of that breakdown group's total across all factors)
- **Statistical testing** - Chi-squared analysis to identify significant patterns
- See also Statistical Significance Testing

Statistical Significance Testing

When you select exactly **one custom column** for breakdown, the factors table includes powerful chi-squared significance testing to identify factors that are preferentially mentioned by different groups.

Show Differences dropdown appears with threshold options:

- **Off** (default)
- **p < .1** (marginally significant)
- **p < .05** (significant)
- **p < .01** (highly significant)
- **p < .005** (very highly significant)

Visual indicators:

- **Significant column** - Shows "Yes" (red highlight) or "No"
- **Cell coloring** - Blue = mentioned more than expected, Orange = mentioned less than expected

Ordinal testing (numeric breakdowns):

- If the chosen breakdown is numeric-like (≥95% of non-missing values parse as numbers), an extra column **Ordinal Sig.** appears.
- It uses the Mantel linear-by-linear association (Cochran–Armitage trend) with ranks 1..k and the same 2×k totals as chi-squared.
- The existing **Significant** column (chi-squared) remains; you can compare both.
- The threshold from Show Differences applies to both tests.

in the factors table when factor-show-differences is on, we calculate chi-sq. but if over 95% of non-missing values in the column selected in #factor-custom-column-input can be interpreted as numeric, we should use an ordinal test instead, or apply an ordinal correction to make the chisq test more powerful

Developer note: Percent mode divides each factor's cell by the group total for that breakdown column. State keys: `factorDisplayMode`, `significanceThreshold`.

Links Panel

🔗 **What you can do here:** View and manage all your causal links in a spreadsheet-like table. You can sort, filter, and edit individual links, view the quotes like a printed page, or export your data to Excel. Each row shows one causal relationship with its source quote and any additional details you've added. Great for detailed review and bulk editing of your causal map data.

Links Table

Links Table Features:

- Standard column filters, sorting , and pagination
- Sentiment column with numeric values (-1 to 1) and blue/white/red conditional formatting (blue = higher, red = lower, white = mid-range relative to the current view)
- **Citation Count** – total number of links in each bundle (cause >> effect pair), with muted green → white conditional formatting (darker green = more links in that bundle relative to the current view)
- **Source Count** – number of different sources contributing links to each bundle, with the same muted green → white conditional formatting
- Checkbox selection for bulk operations
- Edit functionality opens causal overlay for link modification
- Action button to open coding in the Sources pane and scroll to the corresponding highlight
- Clear Table Filters option
- 💡 Tip: For label changes, prefer Factors Search/replace when working on labels across bundles.

Link Editing:

- Single link click opens editor popup
- Multiple link selection opens chooser interface
- Consistent with coding panel behavior

Links Utilities

- Download as Excel
- Take a screenshot and copy it to clipboard
- Clear any filters at the top of the table columns
- Bulk delete any selected rows in the table

Row Grouping and Print View

- **Group by selector** - Choose one or more columns to group rows by values. This applies both the the links table and Print View.

Useful Columns:

- **Bundle** - Shows "cause >> effect" pairs

- **Original Bundle** - If you have used filters which transform the links, like Zoom or Soft Recode, use this to also view the original causes and effects

Print View

The purpose Print View is to make it easy to explore and read actual quotes from the currently filtered links. What it does is show, instead of the contents of the Tabulator table, a printed version of the same information, leaving the table headers and filters in place. The toggle switches between table contents view and print view.

This view prints out the quotes from each row in the table, grouped by the Group By columns formatted as nested headings, and we suppress repeated headings until they change.

We also reveal two more toggles:

- Show Details: Print the values of all the extra columns such as tags and any Custom Columns
- Context: for each quote we add an additional three sentences at each side, highlighting the actual quotes.

You can manually sort the texts using to the sorting widgets in the tabulator headers, as far as allowed by the nested headers.

Search/replace

This works exactly the same as search/replace in the factors table, except that it works on the Cause label and/or the Effect label.

Near the top is a row containing a search box. If you type something into it,

- a replace box and a Replace button also appear.
- the table is filtered to show only matching rows.

The search is **case sensitive**.

You can then alter what you see in the Replace box:

- in the label columns in the table, you see a preview of what the affected rows text so the would look like;
- if you delete all the replace text so it is empty, the preview shows the effect of deleting the search text from each label.

Then when you are satisfied, check all the checkboxes where you want to update the labels as shown. If you want, select all rows using the checkbox at the top of the column. Note, if there are more hits than fit on this page of your table, you'll want to either treat each page separately or increase the page size with the Page Size selector.

Finally, hit the Replace button to actually update the labels as shown in the rows you selected. As you'd expect, this search/replace only affects the factors for the currently selected links: for

example if you have only selected the first three sources, this update will not affect the links in the other sources.

.

Statistics panel

Use this panel to build pivot tables and charts from your project data with a simple drag-and-drop interface (powered by PivotTable.js).

Quick start

1. Choose a data source: Links, Factors (Work in progress!), or Sources.
2. Choose the stage: **After Analysis Pipeline** (matches your filters elsewhere) or **Before Analysis Pipeline** (raw data).
3. Pick a display: Table, Heatmap, or a Plotly chart.

Arrange fields (drag and drop)

- Drag field chips from the left list into the drop zones:
- **Rows**: categories listed down the left.
- **Columns**: categories across the top.
- **Values / Measure**: the numeric field used for calculations (when needed).
- Reorder by dragging within a zone; remove by dragging out or clicking ×.

Choose the calculation ("Summarize by")

- Use the aggregator dropdown to select how numbers are calculated:
- **Count** (default) – how many rows fall into each cell.
- **Sum of / Average of / Min / Max** – pick a numeric field (e.g. `ai_confidence`, text length, counts).
- **Unique Count** – how many distinct values of a field occur.
- Example: to average AI confidence by region and sentiment 1) Put `custom_-Region` in Rows and `sentiment` in Columns 2) Set aggregator to **Average of** and choose `ai_confidence` as the field

Filter or exclude values

- Every field chip has a filter menu. Click it to include/exclude values.
- Use the search box, then tick/untick items. "Select All" toggles everything.
- Filters apply immediately and are remembered in the URL so you can share the view.

Sort and tidy

- Click any row/column header to sort by label or by totals.
- Subtotals and grand totals are included automatically in table views.
- Remove empty columns quickly by excluding the blank value in that field's filter.

Heatmaps and charts

- Switch the renderer to **Heatmap**, **Row Heatmap**, or **Col Heatmap** for quick intensity views (colors match the app palette).

- Choose Plotly renderers for interactive **Bar**, **Line**, **Scatter**, **Stacked Bar**, **Area**, or **Multiple Pie** charts.

Export and sharing

- Use the toolbar to **Copy to Clipboard** or **Download XLSX** of the current table.
- The current configuration is saved in the URL automatically; bookmark or share it with collaborators.

Notes on the datasets

- **Links**: every causal link plus metadata; includes AI fields (e.g. confidence) and reserved columns like `original_cause`, `original_effect`.
- **Factors**: unique factors with frequency, source count, citations, and `original_label` (ALL underlying original labels for the displayed factor, concatenated with line breaks, derived from the current stage's links like `original_cause/effect`).
- **Sources**: document metadata and flattened custom fields (`custom_*`).

💡 Tip: For results that match other panels, use **After Analysis Pipeline**.

Logs panel

The Logs tab shows a concise history of what has happened in your projects – for example, when links or sources were edited or deleted, when projects were created, or when AI processing ran.

- Use the **scope buttons** (This project / All projects) to switch between seeing activity only for the current project or across all projects you have access to.
- Use the **filters above the table** to narrow by user, project name, action type, date, or text in the details column.
- Non-admin users see only activity for projects they own or can edit; admins can see activity across all projects.
- Navigation and other UI clicks are logged only very briefly and are hidden from non-admin users.

Bookmarks and Reports

Report Builder

The Report Builder lets you

- manage your bookmarks
- delete
- copy links
- search and load bookmarks
- create professional reports
    - combining multiple bookmarked views with custom descriptions and variant filtering.
    - with programmatic production of multiple variants e.g. maps with the same filters but for different countries

Requires Pro subscription

📊 **What you can do here:** Build multi-slide reports from your bookmarks. Add markdown descriptions, filter by custom source columns (e.g., country, region), reorder slides, and export everything as formatted HTML ready to paste into Word or Google Docs. Perfect for creating stakeholder reports with multiple views of your data.

Key Features:

- **Drag-and-drop reordering** - Arrange slides in any order
- **Markdown descriptions** - Add rich text titles and explanations with heading styles, lists, and formatting
- **Variant generation** - Create multiple versions of the same map filtered by custom source columns (e.g., one map per country)
- **Include/Exclude toggles** - Selectively include slides in your export
- **HTML export** - Copy formatted HTML with proper heading styles, clickable bookmark links, and legends
- **Persistent settings** - Your slide order, descriptions, and include/exclude states are saved automatically
- Add a bulk delete button to delete all bookmarks with included toggle =ON

Legends:

- For maps we use the full legend
- For tables we use the relevant parts of the legend, omitting material about link and factor colours, sizes and annotation.

Variants:

- Variants simply add an extra filter on top of the bookmark's existing filters (sources + filter pipeline). We do not change any other filters. For example, if the bookmark has 4 selected sources and you choose village=X and Y, then the X variant uses only those of the 4 sources

with village=X, and the Y variant uses only those with village=Y. If a value has no data after the existing filters, no variant is produced.

Where to find the variants: when the accordion section is open, the variants appear as thumbnails below the legend, each preceded by which variable/value it represents (e.g., village: Y). On PDF export and copy to clipboard, first the original bookmark image and its legend are shown, and then each variant with its fullsize image and legend.

How to use:

1. Create bookmarks of your maps and tables
2. Switch to the Report tab (yellow bookmark icon)
3. Click on slide descriptions to edit them (supports markdown: `# Heading`, `## Subheading`, `- List items`)
4. Drag slides to reorder them
5. Use the variant controls to generate multiple versions filtered by custom columns
6. Toggle "Include/Exclude" to control which slides appear in your export
7. Click "Copy HTML" to export all included slides as formatted HTML
8. Paste into Word/Google Docs - headings and links will be preserved

Bookmarks and Slide Decks

Bookmarks

If you want to save a specific view of the app for later, for example an interesting map or table, you can just copy the URL from your browser and paste it somewhere or send to a colleague.

In this section you can learn about the powerful bookmarking and slide decks features.

Requires Pro subscription

📌 **What you can do here:** Save and organize your favorite views of your data. Bookmark specific filter combinations, map layouts, or analysis states so you can quickly return to important insights later. Perfect for preparing presentations or saving different analytical perspectives. From anywhere in the app, find the yellow badge at top right of the screen. Click it! This adds a bookmark to the Bookmarks table.

If there is already a bookmark for this same view, a tick is shown: when a view matches an existing bookmark, show the tick.

When you save a bookmark in the session,

- a small modal, near the badge, opens with a textArea for the Description and Cancel/Save buttons. It is a good idea to provide a description, though it is not required.
- the bookmark is then saved including:
- the screenshot of the currently visible map or table (using the existing handlers for these).
- the current map legend, same as with #copy-legend-btn
- as soon as the screenshot is saved, we open a new small dropdown, headed by the bookmark number, right next to the navbar bookmark badge. This stays open for a few seconds then silently closes but can be opened again on click. This dropdown remains the same, with the same header number, until the next bookmark is recorded. The dropdown has buttons to directly copy to clipboard
- the formatted URL same as #copy-project-link-formatted
- the unformatted URL,
- the screenshot image
- the legend
- **a combination of the bookmark link, the map and the legend**

... and when creating a bookmark on any table tab, when map panel is not showing, remember to also save the screenshot same as for pivot tab #pivot-copy-btn, for links tab #export-links-png-btn , for factors tab #export-factors-png,

The bookmarks table itself is not shown for users below Pro level, i.e. to users only on Free or Private plans.

For Links and Factors tables, bookmarking should record everything: any column reordering and manual column widths, column filtering / sorting...

The Bookmarks Table

- Each row is one bookmark.
- Click the checkbox to select rows.
- Click Load to recall that bookmark, restoring tabs and outputs.
- Click Copy Formatted Link to copy the bookmark's URL to the clipboard.
- Double-click on the Description field to edit it.
- Click Delete to remove the bookmark.
- Click the badge at top right when viewing a map or table to add a new bookmark to it.
- When a new current project is loaded from a URL or by changing the project dropdown, the "Project" filter in the table is pre-filled with the name of the current project.

Main Features:

- **Bookmark badge** - Save/remove current URL state from navbar
- **Bookmarks table** - Manage saved views with sorting and filtering
- **Bulk operations** - Select multiple bookmarks for deletion
- **Editable descriptions** - Double-click to edit bookmark names
- **Auto-normalization** - URLs cleaned and standardized for consistency

Bookmarks Table:

- **Actions** - Load URL, Copy link (plain), Copy formatted link (HTML), Edit URL, Delete individual bookmarks
- **Copy buttons** - Both create short `?bookmark=ID` URLs instead of full parameter strings
- **Columns** - Project, Description (editable), User, URL, Created, ID
- **Bulk delete** - Select multiple with checkboxes, delete with "Bulk:" trash button
- **Badge integration** - Shows bookmark ID when current view is bookmarked

URL Editor:

Click the Edit button next to any bookmark to open the URL Editor, which provides a user-friendly interface to understand complex bookmark parameters. Instead of viewing raw query strings, you see a structured breakdown with proper labels and grouping - showing your selected project, sources, active filters (displayed as individual cards), table settings, and map configurations. The editor categorizes parameters into logical groups (Navigation, Data Selection, Filters & Processing, etc.) and displays filter pipeline details with sequential numbering and status indicators. This makes it easy to understand exactly what state each bookmark preserves without needing technical knowledge of URL parameters.

When a bookmark of a map is saved, the legend text for that map is also saved as part of the `slide_content` JSON and is printed small at the bottom-right of the slide in Slideshow view.

Slide Decks

Coming soon!

🚀 **What you can do here:** Turn your bookmarked views of your data into slides with additional text annotations. Organise your slides into decks. Export your slides as a PDF report. Each project can have one or more decks: collections of bookmarks. Each deck has the same permissions as the project. Each slide is json specifying:

- order (an integer)
- layout (for now just provide two layouts but make it extensible)
- usually a bookmark ID, but may be blank

For now, layouts and themes are hardcoded css snippets. Layouts only deal with position and dimensions, themes only with colours and fonts. Themes and layouts both use the same basic slots or set of components: h1, h2, h3, text1, text2, text3, text4.

in the slide template, remove the subtitle text area. we won't use it any more. Instead, put the Description.

Available layouts:

- Initial slide
- Section slide
- Header and two columns
- Large image on right (2/3 width) and narrower header and text box on the left
- Small image on right (1/3 width) and wider header and text box on the left
- Maximum width/height image and all other components float in front with 90% opacity

In these layouts, the image size is maximised. Text boxes have fixed maximum width and height, and the font size increases virtually up to a sensible maximum to fill the boxes.

The UI is just the existing bookmarks table exposing additional columns for Deck, theme, layout and order and a new action button to open the slide. The slide editor contains the same editable fields plus text fields / textareas for the slots and for bookmark1 (which may be empty), bookmark 2. Plus a toggle to show/hide the bookmark image. Plus a preview of how the slide will look when layout and theme are applied to the components.

A dropdown in the UI lists all decks in the current project plus None. Selecting a deck filters the table to show only bookmarks/slides in that deck. In this mode:

- Newly created bookmarks are added to the selected deck with incremental position (at the end).
- A large sortable Slide Sorter modal lets you reorder slides by dragging (flexbox with 4 slides per row). Each slide has a small vertical widget to duplicate the slide and insert another bookmark/slide ahead of it.

In fullscreen preview mode, small Prev and Next buttons float at mid-left and mid-right, and an Edit button opens the editor for the corresponding slide.

URL state includes a "filter by deck" dropdown. It uses the standard editable selectize to allow creating a new (sanitized, unique) deck name. Bulk actions include "Add selected bookmarks to Current Deck" so you can populate a new deck from existing bookmarks or add new bookmarks directly when a deck is selected. The dropdown lists all decks in the current project plus any newly added names.

Slide previews are slightly debounced while editing text fields in the slide editor.

If the text in any text box (main, sub, text1, text2) exceeds one line, the font size is gradually reduced so a paragraph or two can still fit.

Slides display a very small link to the bookmark (e.g., `?bookmark=nn`) at the bottom-left.

When creating new bookmarks while a slide deck is selected, the new slide uses the same theme and layout as the last slide in that deck.

In the Initial Slide layout, text is placed on a semi-transparent overlay for readability with large image backgrounds.

Styles with dark backgrounds use light fonts. In the "maximum image with overlay text" layout, the overlay uses a dark background, and the Initial Slide overlay also uses a dark background for contrast.

Slide 1 discreetly prints the Causal Map filename, the project's last modified date, and today's date near the bottom.

AI answers panel

Requires an AI subscription

You can open this panel by clicking the corresponding tab lower down on the right-hand side of the app.

🤖 **What you can do here:** Ask questions about your data in plain English and get AI-powered answers. Type questions like "What are the main barriers to education?" and the AI will search through your currently selected sources to provide relevant answers with supporting quotes. Perfect for exploring themes and getting quick insights from large amounts of text.
Main Features:

- **Query input** - Type your questions in plain English
- **Automatic chunking** - Sources split into searchable pieces when needed
- **Similarity slider** - Control search precision (0.1-0.9)
- **Max Chunks slider** - Maximum number of the most relevant chunks to send to the AI
- **Prompt history** - Navigate previous questions with prev/next buttons. See these tips

Search Modes

AI Answers offers two search modes, automatically optimized based on your data size:

Full Sources Mode

Searches the complete text of your sources (documents/interviews).

How it works:

1. Type a question about the text of the currently selected sources
2. System automatically chunks sources into searchable segments (if not already done)
3. Searches through document chunks using AI embeddings and semantic similarity
4. Most relevant chunks are sent to AI for analysis
5. AI generates answers with supporting quotes from your sources

Question expansion and HyDE (Hypothetical Document Embeddings)

Rather than just using the user's question to match against chunks, we call genAI as preparation and ask it to produce:

1. 8 question variants: short phrases likely to appear directly in source texts containing answers
2. 8 answer variants: different longer and shorter phrases which could contain possible answers, substantially and linguistically different from one another

We match each of these phrases against the chosen chunks and make a sum of the scores per chunk, to then select the top n according the the max_chunks slider.

So for example if the user asks what is the connection between money and happiness, the AI produces question variants like:

- having money, being joyful
- being wealthy
- being happy
- connection between money and happiness

And answer variants like:

- financial security enables emotional wellbeing
- wealth contributes to life satisfaction
- economic resources support positive mental health outcomes

Best for: Exploratory questions about raw text, finding themes not yet coded, discovering new patterns.

Link Contexts Mode

Searches only through your coded causal links and their surrounding context (the quote + 3 sentences before/after).

How it works:

1. Gets filtered links from your current filter pipeline (respects Sources dropdown and all Source Groups filters)
2. For each link, extracts the selected quote plus surrounding context
3. Organizes contexts by source, with source metadata (title, custom columns)
4. For ≤500 links: Sends all contexts directly to AI
5. For >500 links: Uses backend semantic search to find most relevant contexts
6. Embeddings generated server-side (via `find-relevant-contexts` edge function)
7. Also uses question expansion (see above)
8. Similarity calculation done server-side using cosine similarity
9. Only relevant context indices returned to frontend
10. No memory/computation overhead in browser
11. AI analyzes contexts showing cause → effect relationships
12. AI uses the cause/effect labels in its narrative (ignoring any original labels if links were recoded)

Context format sent to AI:

```

 
```

Source Interview with Participant 001

ID: ABC-123 custom_Country: Kenya | custom_Gender: Female | custom_Age: 34

Links from this source:

[ABC-123-1] Lack of resources → Poor school performance Context: "We don't have enough books or supplies. The children struggle because..."

[ABC-123-2] Teacher training → Better outcomes Context: "When teachers receive proper training, we see improvements in..." ```

Best for: Questions about causal relationships you've already coded, comparing patterns across sources, analyzing specific demographic groups using Source Groups filters.

Key advantages of Link Contexts mode:

- Uses your coded causal structure, not just raw text
- Respects all your filters (Sources dropdown + Source Groups)
- Includes source metadata in AI context (country, demographics, etc.)
- More focused and structured than full text search
- Automatically scales to large datasets (>500 links) using backend semantic search

etc etc

Settings Panel

⚙️ **What you can do here:**Enable live collaboration. Other settings coming soon.

Live collaboration

Watch as your colleagues make changes to the causal map (and links and factors tables) in real time!

Account Panel

What you can do here: View and manage your personal account settings. Change your password, update your project information, and control your privacy and security settings. This is also where you can export your data or delete your account if needed.

Making projects private requires a Private subscription

User account management and project settings.

Account Features:

- project information and settings
- Password and security management
- Account deletion and data export
- Subscription and billing information

Subscriptions

Subscriptions List

Users without a subscription are either:

- anonymous (not logged in) (this is disabled at present)
- free (logged in)

Subscriptions (via LemonSqueezy) are available in the Account panel.

The subscriptions list uses one row per type (private, pro, team) with seat-count, square radio buttons for monthly/annual and Manual/AI, and a live-updating price (from a JSON price file). Each row includes a text description.

There are three dimensions to the subscriptions,

Manual vs AI,

Type:

- private
- pro
- team

Monthly vs Annual.

User can purchase multiples of one or more subs to distribute to colleagues.

Responses Panel

Work in Progress!

Requires an AI subscription

What you can do here: Review all your AI interactions and usage. See a complete log of AI requests, responses, costs, and performance metrics. Useful for tracking your AI usage, reviewing past queries, and understanding costs. Great for administrators and power users.

Main Features:

- **AI interaction log** - Complete history of all AI requests and responses
- **Cost tracking** - View dollar costs based on token usage and model pricing
- **Performance metrics** - Response times and success rates
- **Sortable table** - Filter and sort by timestamp, model, cost, etc.

Help System

❓ **What you can do here:** Get instant help and guidance while using CausalMap. Search for specific topics, browse documentation by section, or click the question mark icons throughout the app for context-sensitive help. The help system adapts to what you're currently doing. Also view the entire help contents as a separate Guide.

Main Features:

- **Help drawer**: Right-side drawer with collapsible sections
- **Smart search box** at the top of the help drawer. If you enter multiple words, either word matches. To match exact phrases, use quotes.
- **Context-sensitive help**: Click the blue buttons within the app next to open the help drawer.
- **App hints**: When you click a section in the help drawer which refers to part of the app, that element flashes briefly.
- **Standalone help**: The same help information is available as a section within our Guide to causal mapping.

Mobile view

Implemented: Mobile offcanvas for right pane

- A mobile-only two-button toggle (Left side / Right side) sits next to the brand.
- Tapping Right side opens the offcanvas that contains the right pane (tabs and content).
- Tapping Left side closes it. The active side shows a tick.
- The offcanvas shows a "Back to left side" button at the top to close it.
- On desktop (≥ lg), the right pane behaves exactly as before; offcanvas visuals are disabled via CSS.
- A mobile-only Menu consolidates: About, Guide, Help, Bookmark, and Account/Login/Logout.

FAQ - frequently asked questions

Tips on coding

First of all, there's nothing to worry about, it's fun!

The versioning / backups feature means you can always go back to any version of your file at any earlier time point.

Also, bulk editing of factor labels makes it easy to rapidly change one or many links or factors. And you can do it either globally, i.e. changing one factor everywhere in the file, or you can do it for particular sources or specific kinds of links by using filters.

Usually, don't bother coding the same link more than once for the same source, unless they bring up distinctively different evidence each time.

- It's okay not to code a source at all. If there's nothing in it, or if people are just making vague and general sources.
- You'll find you're constantly shifting between sometimes creating new factors, and then going back and reviewing them and merging them and organising them using the bulk editing of factor labels.
- Don't forget you can combine two or more factors into one using the bulk editing of factor labels .
- Don't forget when you want to search rapidly through already coded links through all of the sources, you can click on the rows in the bulk editing of factor labels to go back to the relevant sources directly.
- Occasionally, a source will make a comment about something which is worth coding, even though there isn't actually a causal link. For example, they might make general comments about some outcome without saying what causes it. In this case just use plain coding. (But if you find you are doing this a lot, you might need to rethink your research design.)
- If you are using hierarchical/nested coding (and you probably should) don't forget you can see the whole map zoomed out to the top level: just press the appropriate button in the Filters panel.

Do I have the latest version?

Click the About dropdown at the left of the navigation bar at the top of the app. There you can read the version number. The app should silently update itself when a new version is available.

How can I adjust my links or sources data in bulk (round-tripping)?

As an alternative to editing your links and sources tables manually in the app, you can do what we call round-tripping: download your file, tweak this Excel file manually (e.g. by adding additional columns to the sources tab) and upload it again into a new Causal Map project. Like this you always have to create a new project, which helps you not get mixed up with which version is which.

How can I deal with closed questions like in QuIP?

In Causal Map 4 there is no special treatment of QuIP-style closed questions. What you can do is this:

1) if you want to be able to see the closed question answers while coding, include the answers to the closed questions simply as part of the text of the interview with an appropriate question number. None of this has any meaning to the app, but it might be useful to have for coding. 2) If you want to also use the closed question scores for further analysis, e.g. to make a map of all the interviews which answered an average of better than 0 to some question, then just add a custom column for each question and add the average scores for each question into each column. Then you can apply these values as filters, see here.

What does the 'recycle weakest magnets' slider do?

The slider temporarily removes the N weakest magnets from your list and reassigns their raw labels to the stronger magnets.

For example, if you created 50 magnets but after filtering you only have 5 factors showing with 9% coverage, those 45 weak magnets might be taking evidence away from your main ones. Try moving the slider to 40 to recycle those weakest magnets - this gives their evidence a chance to match with the stronger magnets instead (using the same similarity criterion).

This is most useful when you have lots of fiddly magnets that nibble away at your main ones but then disappear without trace. Note that the slider is a bit unpredictable if you have intervening filters.

