# EXTENDERS

Matienzo Caves Project

*User Manual*

---

*A work-in-progress tool for finding caves, digs and shafts with genuine exploration potential*

## 1. Overview

**Extenders** searches the Matienzo site database and descriptions (5,513 entries) for descriptions that suggest current, unresolved exploration leads. It is designed to help cavers identify promising sites before a trip, or to find targets in a specific area during an expedition.

The tool searches the **description text** of each site record for keywords associated with ongoing exploration — draughts, undescended shafts, continuing passages, digs, and similar features. Results are ranked by how many matching terms appear in the description.

---

## 2. The Search Controls

### 2.1 Area

Select a specific Matienzo area from the dropdown to restrict results to that area, or leave it set to "All Areas" to search the entire database.

The area list is populated from the site index and reflects all 83 areas that appear in the database.

### 2.2 Work Type

Choose one of three modes:

| Mode | Effect |
|---|---|
| **Both** | Searches all descriptions with no mode filter. All keyword panels are shown. |

| Mode | Effect |
|---|---|
| **Surface** | Restricts results to descriptions containing surface-context words (shaft, entrance, shakehole, etc.). Shows only the Surface keyword panel. |
| **Underground** | Restricts results to descriptions containing underground-context words (passage, crawl, chamber, etc.). Shows only the Underground keyword panel. |

### 2.3 Keyword Buttons

Each button represents a category of exploration lead. Clicking a button toggles it on (highlighted) or off. Only active buttons contribute to the search.

**Surface keywords** include: Draughting, Shaft/Pit, Undescended, Surface Dig, Shakehole/Doline, Choked/Blocked, Too Tight, Slot/Small Opening, Loose/Diggable, Resurgence/Sink.

**Underground keywords** include: Draught/Air Flow, Dig/Digging, Unexplored, Undescended Shaft/Aven, Low Crawl Continuing, Holes Between Boulders, Continues/Going, Too Tight (needs enlarging), Aven/Window/High Lead, Active Inlet/Stream, Sump/Dive Lead, Promising/Potential.

Each button matches several synonyms. For example, "Draughting" will match draught, draughting, strong draught, blowing, and air blowing.

**Multiple active buttons are AND conditions** — selecting Draughting and Undescended will only return sites whose descriptions mention both, not either.

#### Quick-select buttons

- **All on** — activates all buttons in the currently visible panel(s)
- **All off** — deactivates all buttons
- **Explorer preset** — activates a recommended set of underground leads: Draught, Dig, Undescended, Low Crawl Continuing, Continues/Going, Too Tight, and Promising/Potential

### 2.4 Additional Terms

A free-text box allows you to add your own search terms, comma-separated. Each term becomes an additional AND condition. For example, entering "calcite, stal" would only return sites mentioning calcite or stal in addition to your selected keywords.

Terms must be at least 3 characters long.

### 2.5 AI Scoring

An optional toggle that sends the top results to the *Claude AI* for scoring and re-ranking. See Section 4 for full details.

---

## 3. Running a Search

### 3.1 Basic workflow

- Select an area (or leave as All Areas)
- Set the Work Type
- Toggle the keyword buttons you want
- Click **Find Extensions**

Results appear below the controls, sorted by keyword hit count. Each result card shows:

- The 4-digit site number, linked to the full site page
- Cave name and area
- A text excerpt showing the relevant part of the description, with matching terms highlighted
- Keyword tags showing which terms matched
- A hit count (number of keyword occurrences in the description)

### 3.2 Interpreting hit counts

A high hit count means the keyword appears many times in the description, which often indicates a site with a long exploration history and multiple leads. It is not always a sign of the most promising site — it may simply be a well-documented cave.

A low hit count from a tight, targeted search (e.g. a specific area with two or three keywords) is often more useful than a high count from a broad search.

### 3.3 Result limit

Results are capped at 200 sites. If your search returns 200, try narrowing it by adding more keywords or selecting a specific area.

### 3.4 Resetting

Click **Reset All** to clear the results and return to a blank search form.

---

## 4. Historical Descriptions and False Positives

This is currently the most important limitation to understand when using Extenders.

### 4.1 The problem

Many site descriptions in the Matienzo database have been updated over decades. A description may begin with an account from the 1980s — "a strongly draughting dig at the end of the passage" — followed by a 2019 update stating that the dig was pushed through and a kilometre of new passage found.

The word **draughting** is still in the description. Extenders will match it. But the lead is long gone.

The same issue applies to:

- Passages described as "too tight" that were subsequently enlarged
- Chokes described as blocked that were later dug through
- Undescended pitches that have since been rigged and explored
- Digs mentioned in historical accounts that were completed years ago

### 4.2 What Extenders does to mitigate this

Two mechanisms are in place to reduce — but not eliminate — false positives:

- **Sentence-level lead scoring:** Each description is split into sentences. A sentence containing an exploration keyword alongside a "resolved" phrase (was passed, has been pushed, been dug through, now open, etc.) is flagged as a resolved lead rather than a current one. Sites where resolved leads clearly outnumber open leads are shown with a dimmed card and a warning.
- **Position-based undescended filter:** For surface searches, the word "undescended" is only accepted as a current lead if it appears in the opening body of the description (the first 400 characters after the header). A large cave that merely contains an undescended pitch somewhere inside it will not appear in surface undescended searches.

### 4.3 What to do

Treat Extenders results as a starting point, not a definitive list. The excerpt shown on each card is the most relevant passage found — read it critically. If it reads as historical narrative, click through to the full site page to check recent updates.

> **Note:** *Cards with a dimmed grey border carry an automatic warning that leads in that description may have been passed. These sites are worth a quick check but are lower priority.*

---

## 5. AI Scoring

### 5.1 What it does

When AI scoring is enabled and you run a search, the top results are sent to the *Claude AI* (*Haiku model*). Claude reads each excerpt and assigns a score of 1 to 10 for exploration potential, along with a brief 12-word reason. The results are then re-sorted by AI score.

The AI understands context in a way that keyword matching cannot — it can recognise that "the draught has since stopped" is a less promising lead than "blowing strongly outward, too tight, needs enlarging."

### 5.2 Recommended workflow

AI scoring works best as a second pass, not a first one:

| | |
|---|---|
| **1** | Set your area, work type and keywords. Leave AI scoring switched **off**. |
| **2** | Click **Find Extensions**. Review the results. If the number and quality look useful, proceed. |
| **3** | Switch AI scoring **on**, then click **Find Extensions** again. The same search runs, and this time the results are scored and re-ranked by the AI. |
| **4** | Each card now shows an AI score (1-10) and a brief reason. Sites with a green border are ranked highly by the AI. |

> **Note:** *AI scoring only re-ranks the results returned by the keyword search. It does not find additional sites. Run your keyword search first to make sure it is returning a sensible set before applying AI scoring.*

### 5.3 Limitations

The AI scores each site based on the excerpt text only (approximately 400 characters). It does not have access to the full description, survey data, or the site's exploration history. Scores are indicative rather than definitive.

AI scoring adds a few seconds to the search time and incurs a small API cost. For occasional use this is negligible.

---

## 6. Result Card Colours

| Border colour | Meaning |
|---|---|
| **Gold (left border)** | High keyword hit count — the search terms appear frequently in the description. |
| **Green (left border)** | High AI score (8 or above out of 10) — requires AI scoring to be active. |

| Border colour | Meaning |
|---|---|
| **Grey / dimmed** | Leads in this description appear to have been passed or resolved. Check the full site page before adding to your list. |
| **No highlight** | Standard result — keyword matches found but no special flags. |

---

## 7. Practical Tips

- **Start specific.** A search for one area with two or three tightly chosen keywords is more useful than an all-areas search with many keywords.
- **AND logic narrows results.** Each additional active keyword button reduces the result count, because all selected keywords must appear in the same description.
- **Use the excerpt.** The highlighted excerpt shows the most relevant passage in the description. If it reads as historical or resolved, use the site number link to check the full page.
- **Grey cards are not useless.** A dimmed card may still contain live leads that were not caught by the resolved-lead detection. Always read the excerpt.
- **AI scoring is a second opinion, not a filter.** Run the keyword search first, check the result count, then apply AI scoring if the results look promising.
- **Reset All clears everything.** Including the results list. If you want to refine a search, change the controls and click Find Extensions again — do not reset unless you want a completely fresh start.

---

*Matienzo Caves Project — Extenders is a work in progress. Feedback welcome.*