Pattern Extractor

Simple offline tool for extracting variable-length substrings (1-20 chars) from pasted text using a 20-position Wordle-style pattern row.

Quick start

• Paste your source text into the Input box.

• Define your pattern in the 20 squares: choose a type above each square and optionally enter a fixed character in the square.

• Leave Length on Auto, or set a manual Length (1-20).

• Click Extract.

• Use Output Options to format the results, then Copy Output.

What each area does

Area	Purpose
Pattern (20 positions)	Defines which substrings to extract. Each position has a Type selector and an optional fixed character.
Length Auto / Length (1-20)	Controls extraction length L. Auto computes L from your pattern; manual forces L.
Input	Paste any text. The app scans it as-is (including spaces, punctuation, newlines).
Output Options	Controls quoting, dedupe, trim-for-output, and output layout (single line vs CSV row).
Output	Shows the formatted list of extracted matches. Use Copy Output to copy to clipboard.

Pattern row rules

The pattern row always shows 20 positions. Each position has:

• A Type selector: L (Letter), N (Number), C (Character), * (Wildcard).

• A square input: optional fixed character for that position.

Type selectors

Code	Meaning	What matches
L	Letter	A-Z (case-insensitive)
N	Number	0-9
C	Character	Any single character, including space
*	Wildcard	Any single character (same as C, separate for UX)

Fixed characters

If you type a character into a square, the extracted substring must have that exact character at that position. For L (Letter), matching is case-insensitive; for other types, matching is exact.

If a square is blank, any character consistent with the chosen type is allowed at that position.

Invalid patterns

If you enter a character that conflicts with the selected type (for example, you enter 'A' but the type is N), that square is highlighted as invalid and extraction is blocked until fixed.

Length (1-20) logic

The app extracts substrings of length L. You can run in Auto length or Manual length mode.

Auto length

When Length Auto is enabled, L is computed from your pattern. L is the greatest position (1-based) where either:

• A square contains a character, OR

• The type selector is not * (Wildcard).

Auto length always uses at least L=1.

Manual length

Disable Length Auto to set Length (1-20) explicitly. Only positions 1..L participate in matching and extraction; positions beyond L are ignored and visually dimmed.

How matching works

The app scans your input text as a single continuous string, without removing or normalizing characters. It checks every possible substring of length L and keeps those that satisfy all per-position rules.

This includes matches across punctuation, tabs, and newlines (because those are just characters in the text).

Uniqueness and trimming

By default, duplicates are removed (unique output). Enable Allow duplicates to keep repeated matches. Trim matches removes leading/trailing whitespace for output only; matching always uses the raw substring.

Output formatting

You can format output for quick paste into SQL, Excel, or Power BI.

Wrap each in single quotes

When enabled, each match becomes 'value' before it is joined.

Layout modes

• Single line: outputs one comma-separated line with no newlines.

• Single row (CSV): outputs CSV-like content. If Header is ON, the first row is the header name and the second row is the values.

CSV quoting rules

In Single row (CSV) mode, values are quoted with double quotes only when needed (comma, newline, or double quote present). Embedded double quotes are escaped by doubling them.

Buttons and counters

• Extract: runs the scan and writes formatted results to Output.

• Copy Output: copies output to clipboard (some browsers restrict clipboard on file://; in that case the output is selected for Ctrl/Cmd+C).

• Clear Input / Clear Output: clears the text areas.

• Reset Pattern: resets all types to * and clears all squares; sets length back to Auto.

• Sample Input: loads a demo pattern and demo text so you can see the tool working.

Counters show: Total matches (before dedupe), Unique output count, Time taken (ms), and Active length L.

Troubleshooting

• **No matches found:** Try loosening the pattern: remove fixed characters, switch some positions to * or C, or reduce L.

• **Cannot extract - invalid pattern:** One or more squares contains a character that conflicts with its type selector. Fix or clear the square.

• **Clipboard unavailable:** Some browsers block clipboard access for local files. Use Copy Output, then press Ctrl/Cmd+C after the output is selected.

• **Too many results:** Disable Allow duplicates (default), increase specificity with fixed characters, or increase L to reduce accidental matches.

Privacy and offline use

This is a static offline page. It does not upload data anywhere. The only way data leaves the page is if you copy/paste it into another application.