

## Hakrawler Professional Manual

### Introduction

Hakrawler is a fast and lightweight web crawler built for reconnaissance. Designed by hakluke, it is tailored to uncover hidden endpoints, JavaScript files, forms, and internal links from web applications — making it an essential tool in the arsenal of ethical hackers, penetration testers, and bug bounty hunters.

Unlike traditional crawlers, Hakrawler emphasizes speed, security relevance, and integration with other reconnaissance tools. This guide provides you with step-by-step instructions and real-world examples using the domain <https://cust.edu.pk> as a sample target.

### Installation Steps (Kali Linux or Debian-based)

#### 1. Install Go Language (if not installed):

```
sudo apt update
sudo apt install golang -y
```

**Objective:** Install the Go programming language, required to compile Hakrawler.

```
(naqash@kali)-[~]
$ sudo apt update
sudo apt install golang -y

Hit:1 http://kali.org/kali kali-rolling InRelease
Hit:2 https://repo.mongodb.org/apt/debian bookworm/mongodb-org/7.0 InRelease
237 packages can be upgraded. Run 'apt list --upgradable' to see them.
Warning: https://repo.mongodb.org/apt/debian/dists/bookworm/mongodb-org/7.0/InRelease: Policy will reject signature within a year, see --audit for details
Installing:
golang

Installing dependencies:
golang-1.24 golang-1.24-doc golang-doc

Summary:
Upgrading: 0, Installing: 4, Removing: 0, Not Upgrading: 237
Download size: 138 kB
Space needed: 713 kB / 27.3 GB available

Get:1 http://kali.download/kali kali-rolling/main amd64 golang-1.24-doc all 1.24.4-1 [112 kB]
Get:2 http://kali.download/kali kali-rolling/main amd64 golang-1.24 all 1.24.4-1 [16.2 kB]
Get:3 http://kali.org/kali kali-rolling/main amd64 golang-doc all 2:1.24-2 [5,148 B]
Get:4 http://kali.org/kali kali-rolling/main amd64 golang amd64 2:1.24-2 [5,088 B]
Fetched 138 kB in 47s (2,941 B/s)
Selecting previously unselected package golang-1.24-doc.
(Reading database ... 500054 files and directories currently installed.)
Preparing to unpack .../golang-1.24-doc_1.24.4-1_all.deb ...
Unpacking golang-1.24-doc (1.24.4-1) ...
Selecting previously unselected package golang-1.24.
Preparing to unpack .../golang-1.24_1.24.4-1_all.deb ...
Unpacking golang-1.24 (1.24.4-1) ...
Selecting previously unselected package golang-doc.
Preparing to unpack .../golang-doc_2:1.24-2_all.deb ...
Unpacking golang-doc (2:1.24-2) ...
Selecting previously unselected package golang:amd64.
Preparing to unpack .../golang_2:1.24-2_amd64.deb ...
Unpacking golang:amd64 (2:1.24-2) ...
Setting up golang-1.24-doc (1.24.4-1) ...
Setting up golang-1.24 (1.24.4-1) ...
```

## 2. Install Hakrawler using Go:

go install [github.com/hakluke/hakrawler@latest](https://github.com/hakluke/hakrawler)

**Objective:** Download and install the latest version of Hakrawler.

```
(naqash@kali)-[~]
$ go install github.com/hakluke/hakrawler@latest

go: downloading github.com/hakluke/hakrawler v0.0.0-20241221204001-76152559a809
go: downloading github.com/gocolly/colly/v2 v2.1.1-0.20220308084714-a61109486557
go: downloading github.com/antchfx/xmlquery v1.3.9
go: downloading github.com/antchfx/htmlquery v1.2.4
go: downloading github.com/PuerkitoBio/goquery v1.8.0
go: downloading github.com/nlnwa/whatwg-url v0.1.0
go: downloading github.com/temoto/robotstxt v1.1.2
go: downloading golang.org/x/net v0.0.0-20220225172249-27dd8689420f
go: downloading google.golang.org/appengine v1.6.7
go: downloading github.com/antchfx/xpath v1.2.0
go: downloading github.com/golang/groupcache v0.0.0-20210331224755-41bb18fe9da
go: downloading github.com/andybalholm/cascadia v1.3.1
go: downloading github.com/willf/bitset v1.1.10
go: downloading golang.org/x/text v0.3.7
go: downloading google.golang.org/protobuf v1.27.1
```

## 3. Add Hakrawler to your system PATH:

export PATH=\$PATH:\$(go env GOPATH)/bin

**Objective:** Ensure Hakrawler is accessible from any terminal session.

```
(naqash@kali)-[~]
$ export PATH=$PATH:$(go env GOPATH)/bin cat {} \; | go run -link
```

## 4. Verify Installation:

hakrawler -h

**Objective:** Confirm that Hakrawler is installed and working properly.

```
(naqash@kali)-[~]
$ hakrawler -h
Usage of hakrawler:
  -d int          Depth to crawl. (default 2)
  -dr            Disable following HTTP redirects.
  -h string       Custom headers separated by two semi-colons. E.g. -h "Cookie: foo=bar;;Referer: http://example.com/"
  -i            Only crawl inside path
  -insecure       Disable TLS verification.
  -json          Output as JSON.
  -proxy string   Proxy URL. E.g. -proxy http://127.0.0.1:8080
  -s            Show the source of URL based on where it was found. E.g. href, form, script, etc.
  -size int       Page size limit, in KB. (default -1)
  -subs          Include subdomains for crawling.
  -t int         Number of threads to utilise. (default 8)
  -timeout int    Maximum time to crawl each URL from stdin, in seconds. (default -1)
  -u            Show only unique urls.
  -w            Show at which link the URL is found.
```

## 5. Basic crawl of homepage:

echo "https://www.cust.edu.pk" | hakrawler

**Objective:** Perform a default-depth crawl to discover URLs on the homepage.

```
naqash@kali:~$ echo "https://www.cust.edu.pk" | hakrawler
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/pnd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
```

## Core Flags

Flag	Description
-d [num]	Crawl depth (default: 2)
-u	Display only unique URLs
-subs	Include subdomains (e.g., assets.cust.edu.pk)
-json	Output in JSON format
-s	Show where each URL was discovered (href, script, etc.)
-w	Show parent URL where link was found
-t [num]	Number of concurrent threads (default: 8)
-timeout [s]	Per-URL timeout in seconds
-h	Add custom headers (format: "Key: value;;Key2: value2")
-proxy	Use HTTP/SOCKS proxy
-insecure	Disable SSL/TLS verification

## 6. Increase crawl depth to 3:

```
echo "https://www.cust.edu.pk" | hakrawler -d 3
```

**Objective:** Discover deeper pages by crawling 3 levels into the site.

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -d 3 -u
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/bs-english/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-mathematics/phil-mathematics/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/ms-biosciences/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/ms-civil-engineering-evening/
```

## 7. Show only unique URLs:

```
echo "https://www.cust.edu.pk" | hakrawler -u
```

**Objective:** Filter and output only unique links (remove duplicates).

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -u
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-mathematics/phil-mathematics/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/ms-biosciences/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/ms-civil-engineering-evening/
```



8. Include subdomains during crawl:

```
echo "https://www.cust.edu.pk" | hakrawler -subs
```

**Objective:** Include URLs from subdomains like student.cust.edu.pk or library.cust.edu.pk.

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -subs

https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/bs-english/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
```

9. Show where each link was found:

```
echo "https://www.cust.edu.pk" | hakrawler -s
```

**Objective:** Identify the source of each discovered URL (e.g., script, href, form).

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -s

[href] https://cust.edu.pk/
[href] https://cust.edu.pk/
[href] https://cust.edu.pk/
[href] https://cust.edu.pk/about/
[href] https://cust.edu.pk/about-capital/messages/chancellors-message/
[href] https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
[href] https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
[href] https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
[href] https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
[href] https://cust.edu.pk/my-capital/allpolicy/general-policies/
[href] https://cust.edu.pk/my-capital/allpolicy/students-discipline/
[href] https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
[href] https://cust.edu.pk/my-capital/academic-calendar/
[href] https://cust.edu.pk/scholarship/
[href] https://pay.cust.edu.pk/
[href] https://www.cust.edu.pk/placement-office/
[href] https://www.cust.edu.pk/green-cust/
[href] https://www.cust.edu.pk/sustainability/
[href] https://www.cust.edu.pk/about-gym/
[href] https://www.cust.edu.pk/my-capital/vis-directorate/
[href] https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
[href] https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
[href] https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
[href] https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
[href] https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
[href] https://www.cust.edu.pk/bs-cyber-security/
[href] https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
[href] https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
[href] https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
[href] https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
[href] https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
[href] https://cust.edu.pk/faculty-of-management-and-social-sciences/bs-english/
[href] https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
[href] https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
[href] https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
[href] https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
[href] https://cust.edu.pk/faculty-of-computing/department-of-mathematics/mphil-mathematics/
```

## 10. Output results in JSON format:

echo "https://www.cust.edu.pk" | hakrawler -json

**Objective:** Produce machine-readable output for scripts or automation.

```
(naqash@kali)~$ echo "https://www.cust.edu.pk" | hakrawler -json
{"Source":"href","URL":"https://cust.edu.pk/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/about/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/about-capital/messages/chancellors-message/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/ms-regulations/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/phd-regulations/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/general-policies/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/students-discipline/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/allpolicy/university-dress-code/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/my-capital/academic-calendar/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/scholarship/","Where":""}
{"Source":"href","URL":"https://pay.cust.edu.pk/","Where":""}
{"Source":"href","URL":"https://www.cust.edu.pk/placement-office/","Where":""}
{"Source":"href","URL":"https://www.cust.edu.pk/green-cust/","Where":""}
{"Source":"href","URL":"https://www.cust.edu.pk/sustainability/","Where":""}
{"Source":"href","URL":"https://www.cust.edu.pk/about-gym/","Where":""}
{"Source":"href","URL":"https://www.cust.edu.pk/my-capital/vis-directorate/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-management-sciences/bachelor-of-business-administration-bba/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-engineering/departement-of-electrical-and-computer-engineering/bs-electrical-engineering/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/bs-computer-science/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/bs-software-engineering/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-software-engineering-3/bs-artificial-intelligence-2/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-engineering/departement-of-civil-engineering/bs-civil-engineering/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-engineering/departement-of-mechanical-engineering/bs-mechanical-engineering/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-health-and-life-sciences/departement-of-bioinformatics-and-bio-sciences/bs-biotechnology/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-accounting-and-finance/bs-accounting-and-finance/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-psychology/bs-psychology/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-pharmacy/departement-of-pharmacy/pharm-d/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/ms-computer-science/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-software-engineering-3/ms-artificial-intelligence-2/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-engineering/departement-of-electrical-and-computer-engineering/ms-electrical-engineering/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-computing/departement-of-mathematics/phil-mathematics/","Where":""}
{"Source":"href","URL":"https://cust.edu.pk/faculty-of-health-and-life-sciences/departement-of-bioinformatics-and-bio-sciences/ms-biosciences/","Where":""}
```

## 11. Show source of each link:

echo "https://www.cust.edu.pk" | hakrawler -s -w

```
(naqash@kali)~$ echo "https://www.cust.edu.pk" | hakrawler -s -w
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/about/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/about-capital/messages/chancellors-message/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/general-policies/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/students-discipline/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/my-capital/academic-calendar/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/scholarship/
[https://www.cust.edu.pk/] [href] https://pay.cust.edu.pk/
[https://www.cust.edu.pk/] [href] https://www.cust.edu.pk/placement-office/
[https://www.cust.edu.pk/] [href] https://www.cust.edu.pk/green-cust/
[https://www.cust.edu.pk/] [href] https://www.cust.edu.pk/sustainability/
[https://www.cust.edu.pk/] [href] https://www.cust.edu.pk/about-gym/
[https://www.cust.edu.pk/] [href] https://www.cust.edu.pk/my-capital/vis-directorate/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-management-sciences/bachelor-of-business-administration-bba/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-engineering/departement-of-electrical-and-computer-engineering/bs-electrical-engineering/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/bs-computer-science/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/bs-software-engineering/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-computing/departement-of-software-engineering-3/bs-artificial-intelligence-2/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-engineering/departement-of-civil-engineering/bs-civil-engineering/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-engineering/departement-of-mechanical-engineering/bs-mechanical-engineering/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-health-and-life-sciences/departement-of-bioinformatics-and-bio-sciences/bs-biotechnology/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-accounting-and-finance/bs-accounting-and-finance/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-management-and-social-sciences/departement-of-psychology/bs-psychology/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-pharmacy/departement-of-pharmacy/pharm-d/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-computing/departement-of-computer-science/ms-computer-science/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-computing/departement-of-software-engineering-3/ms-artificial-intelligence-2/
[https://www.cust.edu.pk/] [href] https://cust.edu.pk/faculty-of-engineering/departement-of-electrical-and-computer-engineering/ms-electrical-engineering/
```



## 12. Custom User-Agent header:

echo "https://www.cust.edu.pk" | hakrawler -h "User-Agent: ReconScanner/1.0"

```
(naqash@kali)-[~]
$ echo "https://www.cust.edu.pk" | hakrawler -h "User-Agent: ReconScanner/1.0"

https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
```

## 13. Use custom headers (like fake User-Agent):

echo "https://www.cust.edu.pk" | hakrawler -h "User-Agent: Mozilla/5.0;;Referer: https://google.com"

**Objective:** Bypass simple bot protections and mimic real browser requests.

```
(naqash@kali)-[~]
$ echo "https://www.cust.edu.pk" | hakrawler -h "User-Agent: Mozilla/5.0;;Referer: https://google.com"

https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/ms-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-mathematics/phil-mathematics/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/ms-biosciences/
```

14. Set request timeout to 10 seconds:

```
echo "https://www.cust.edu.pk" | hakrawler -timeout 10
```

**Objective:** Limit how long to wait for each response to speed up crawling.

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -timeout 10
https://cust.edu.pk/
https://cust.edu.pk/
https://cust.edu.pk/about/
https://cust.edu.pk/about-capital/messages/chancellors-message/
https://cust.edu.pk/about-capital/messages/vice-chancellor-s-message/
https://cust.edu.pk/my-capital/allpolicy/undergrad-regulations/
https://cust.edu.pk/my-capital/allpolicy/phd-regulations/
https://cust.edu.pk/my-capital/allpolicy/general-policies/
https://cust.edu.pk/my-capital/allpolicy/students-discipline/
https://cust.edu.pk/my-capital/allpolicy/university-dress-code/
https://cust.edu.pk/my-capital/academic-calendar/
https://cust.edu.pk/scholarship/
https://pay.cust.edu.pk/
https://www.cust.edu.pk/placement-office/
https://www.cust.edu.pk/green-cust/
https://www.cust.edu.pk/sustainability/
https://www.cust.edu.pk/about-gym/
https://www.cust.edu.pk/my-capital/vis-directorate/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-management-sciences/bachelor-of-business-administration-bba/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/bs-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/bs-software-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/bs-artificial-intelligence-2/
https://www.cust.edu.pk/bs-cyber-security/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/bs-civil-engineering/
https://cust.edu.pk/faculty-of-engineering/department-of-mechanical-engineering/bs-mechanical-engineering/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/bs-biotechnology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-accounting-and-finance/bs-accounting-and-finance/
https://cust.edu.pk/faculty-of-management-and-social-sciences/department-of-psychology/bs-psychology/
https://cust.edu.pk/faculty-of-management-and-social-sciences/bs-english/
https://cust.edu.pk/faculty-of-pharmacy/department-of-pharmacy/pharm-d/
https://cust.edu.pk/faculty-of-computing/department-of-computer-science/ms-computer-science/
https://cust.edu.pk/faculty-of-computing/department-of-software-engineering-3/ms-artificial-intelligence-2/
https://cust.edu.pk/faculty-of-engineering/department-of-electrical-and-computer-engineering/ms-electrical-engineering/
https://cust.edu.pk/faculty-of-computing/department-of-mathematics/phil-mathematics/
https://cust.edu.pk/faculty-of-health-and-life-sciences/department-of-bioinformatics-and-bio-sciences/ms-biosciences/
https://cust.edu.pk/faculty-of-engineering/department-of-civil-engineering/ms-civil-engineering-evening/
```

15. Save output to file:

```
echo "https://www.cust.edu.pk" | hakrawler -u > cust-urls.txt
```

**Objective:** Store discovered URLs in a file for later analysis.

```
(naqash@kali)-[~]
└─$ echo "https://www.cust.edu.pk" | hakrawler -u > cust-urls.txt
(naqash@kali)-[~]
```