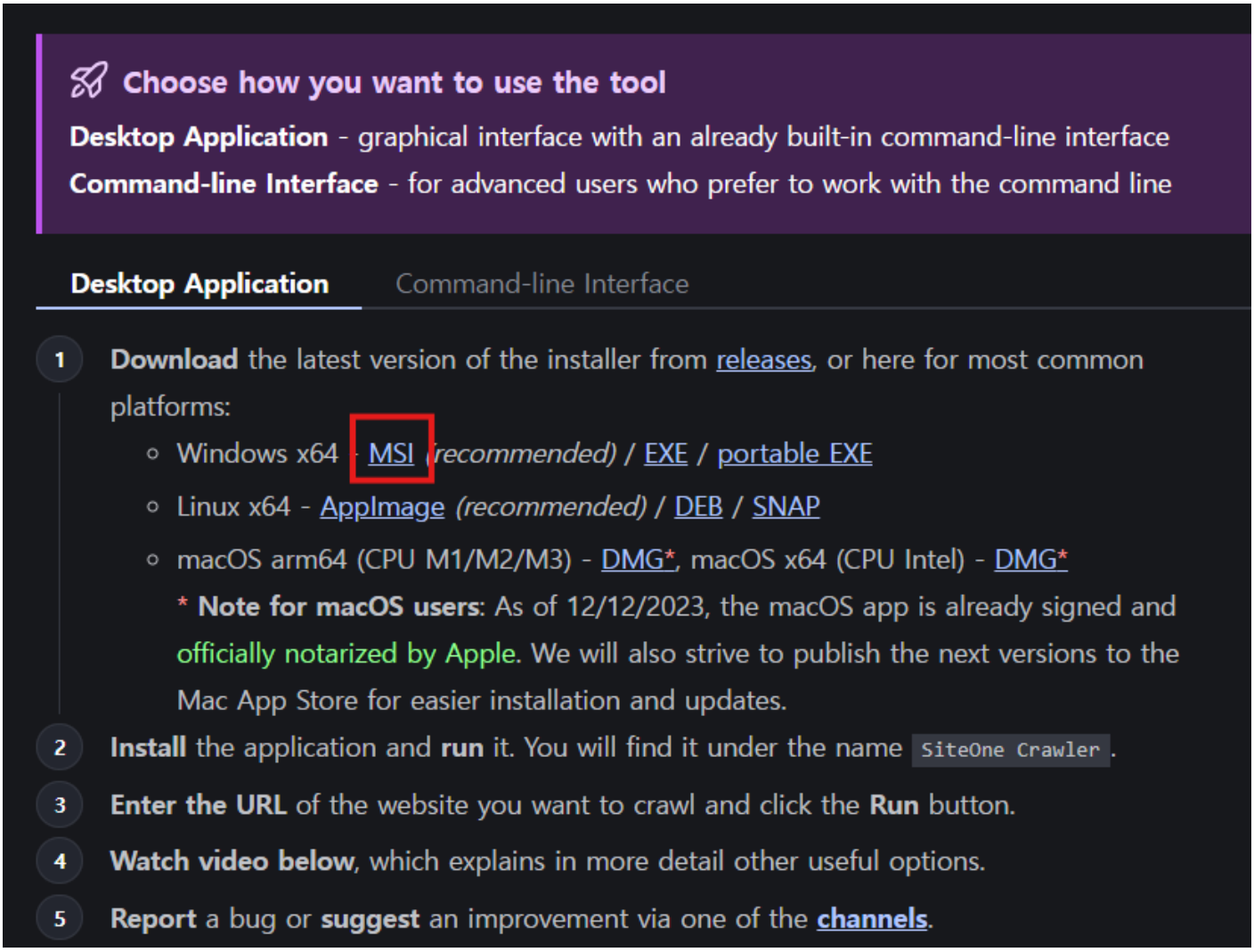
다보리- 기타

- 크롤링 툴을 이용한 h1 태그 없는 페이지 찾기
- 크롤링 툴을 이용한 h1 2개 이상인 페이지 찾기



크롤링 툴을 이용한 h1 태그 없는 페이지 찾기

1. 링크를 클릭합니다.

2. 아래 이미지와 같이 MSI 링크를 클릭하여 crawl 툴을 설치합니다.

Choose how you want to use the tool

Desktop Application - graphical interface with an already built-in command-line interface

Command-line Interface - for advanced users who prefer to work with the command line

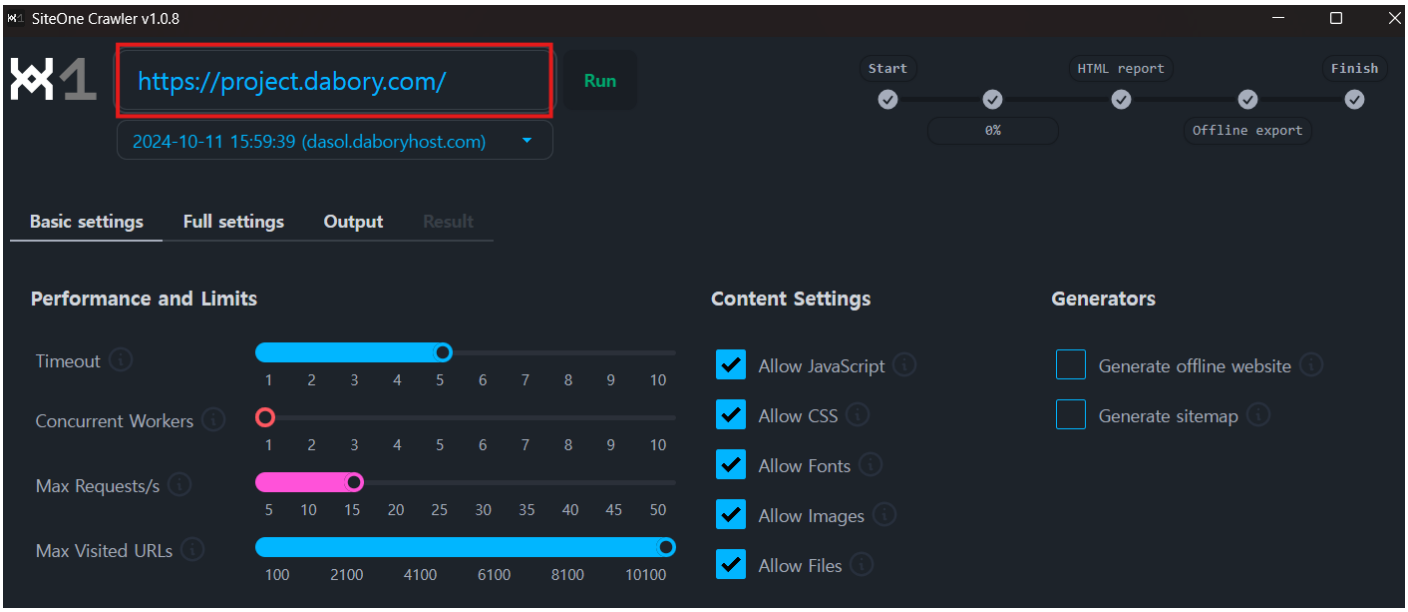
Desktop Application
Command-line Interface

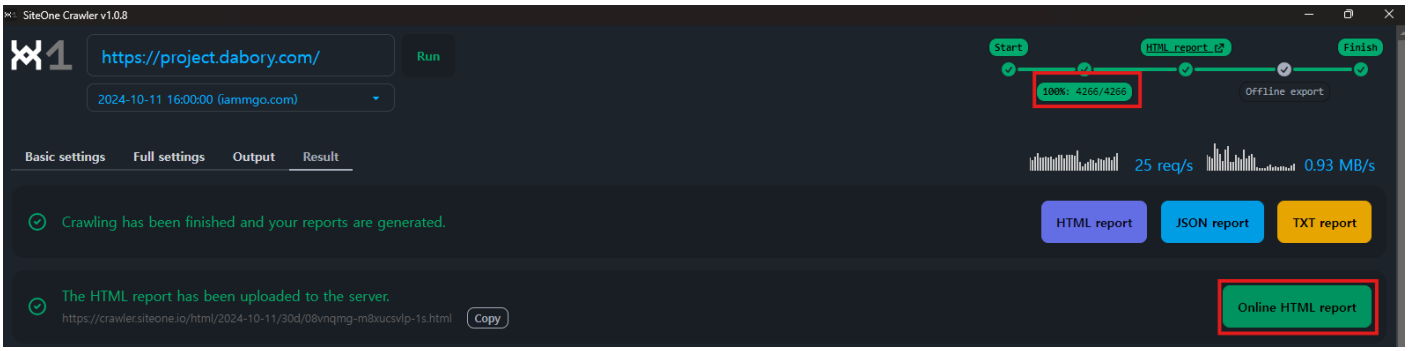
1. **Download** the latest version of the installer from releases, or here for most common platforms:
    - Windows x64 - MSI *(recommended)* / EXE / portable EXE
    - Linux x64 - AppImage *(recommended)* / DEB / SNAP
    - macOS arm64 (CPU M1/M2/M3) - DMG*, macOS x64 (CPU Intel) - DMG*

      * **Note for macOS users**: As of 12/12/2023, the macOS app is already signed and officially notarized by Apple. We will also strive to publish the next versions to the Mac App Store for easier installation and updates.
2. **Install** the application and **run** it. You will find it under the name `SiteOne Crawler`.
3. **Enter the URL** of the website you want to crawl and click the **Run** button.
4. **Watch video below**, which explains in more detail other useful options.
5. **Report** a bug or **suggest** an improvement via one of the **channels**.

3. 설치 완료후 툴을 실행합니다.

4. 아래 이미지와 같이 붉은 박스에 크롤링할 사이트의 URL을 입력합니다.





SiteOne Crawler v1.0.8
Run
Start
HTML report
Finish
0%
Offline export
Basic settings
Full settings
Output
Result
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10
5 10 15 20 25 30 35 40 45 50
100 2100 4100 6100 8100 10100

https://project.dabory.com/

2024-10-11 15:59:39 (dasol.daboryhost.com)

Performance and Limits

Timeout

Concurrent Workers

Max Requests/s

Max Visited URLs

Content Settings

☑ Allow JavaScript

☑ Allow CSS

☑ Allow Fonts

☑ Allow Images

☑ Allow Files

Generators

☐ Generate offline website

☐ Generate sitemap

5. URL 입력창 우측에 Run 버튼을 클릭합니다.

6. 아래 이미지와 같이 100% 완료가 되면 Online HTML report 버튼을 클릭합니다.

SiteOne Crawler v1.0.8
Run
Start
HTML report
Finish
Offline export
Basic settings
Full settings
Output
Result
HTML report
JSON report
TXT report
Copy
Online HTML report

https://project.dabory.com/

2024-10-11 16:00:00 (iammgo.com)

100%: 4266/4266

25 req/s

0.93 MB/s

Crawling has been finished and your reports are generated.

The HTML report has been uploaded to the server.

https://crawler.siteone.io/html/2024-10-11/30d/08vnqmg-m8xucsvlp-1s.html

7. 아래 이미지와 같이 SEO and Open Graph 메뉴를 클릭합니다.

Summary 6 12 5 9 3
Visited URLs 526 208 3532
Best practices 4 2841 2 2013
Accessibility 920 2750 1128
Security 3346 4800 960 1920
SEO and OpenGraph 1 1 1
Heading structure 256 457
Image Gallery 1143
404 URLs 524
Redirected URLs 208

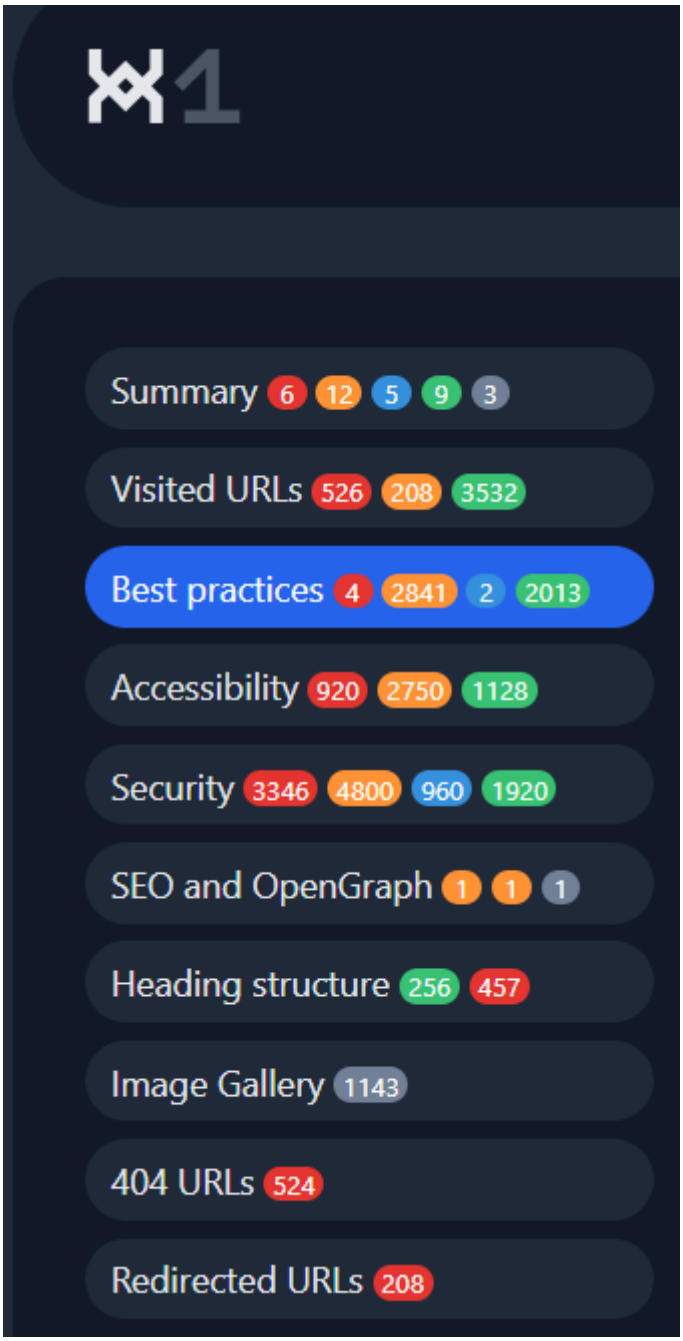
8. 아래 이미지와 같이 H1으로 정렬하여 H1 태그가 없는 페이지를 확인합니다.

SEO metadata

Fulltext search

Found 713 row(s).

URL	Indexing	Title	H1	Description	Keywords
/ai-chatbot-solutions	Allowed		Missing H1		
/ai-matching-solutions	Allowed		Missing H1		
/bbs/list/elastic-functions	Allowed		Missing H1		
/bbs/list/elastic-requirements	Allowed		Missing H1		
/bbs/list/elastic-showcases	Allowed		Missing H1		
/bbs/list/erp-functions	Allowed		Missing H1		
/bbs/list/erp-requirements	Allowed		Missing H1		
/bbs/list/pos-functions	Allowed		Missing H1		
/bbs/list/pos-requirements	Allowed		Missing H1		
/bbs/list/pos-showcases	Allowed		Missing H1		
/bbs/list/shop-functions	Allowed		Missing H1		
/bbs/list/sso-showcases	Allowed		Missing H1		
/maint-matching	Allowed		Missing H1		
/partner-addnew	Allowed		Missing H1		
/partner-coexistence	Allowed		Missing H1		
/partner-showcases	Allowed		Missing H1		
/platform-showcases	Allowed		Missing H1		
/bbs/list/asp-showcases	Allowed		ASP 유지보수 사례		
/bbs/list/asp-showcases?page=1	Allowed		ASP 유지보수 사례		



크롤링 툴을 이용한 h1 2개 이상인 페이지 찾기

1. 아래 이미지와 같이 Best practices 메뉴를 클릭합니다.

Summary 6 12 5 9 3
Visited URLs 526 208 3532
Best practices 4 2841 2 2013
Accessibility 920 2750 1128
Security 3346 4800 960 1920
SEO and OpenGraph 1 1 1
Heading structure 256 457
Image Gallery 1143
404 URLs 524
Redirected URLs 208

2. 아래 이미지와 같이 각각의 url을 클릭하여 페이지를 확인한뒤 수정합니다.

Severity	Occurs	Detail	Affected URLs (max 5)
critical	488	No <h1> tag found in the HTML content.	URL 1, URL 2, URL 3, URL 4, URL 5
critical	5	Multiple <h1> headings found.	URL 1, URL 2, URL 3, URL 4
warning	451	Heading structure is skipping levels: found an <h3> after an <h1>.	URL 1, URL 2, URL 3, URL 4, URL 5
warning	232	Heading structure is skipping levels: found an <h4> after an <h1>.	URL 1, URL 2, URL 3, URL 4, URL 5
warning	14	Heading structure is skipping levels: found an <h6> after an <h1>.	URL 1, URL 2, URL 3, URL 4, URL 5
warning	1	Heading structure is skipping levels: found an <h5> after an <h1>.	/sso-solutions