Static HTML sites and Exponential

Moving a hand-made or generated static HTML site into Exponential 6 and back - reading the files or crawling the site, extracting the content area, HTML to XmlText, images and documents, keeping .html addresses, and...

Source file: `doc/eng-us/install/migration/static-html.md` Last changed 2026-11-04
https://doc.exponential.earth/learninminutes/exphelp/viewer/install/migration/static-html.md

A static site is the simplest source: every page is a file you can read. The work is in separating the content from the repeated layout around it, and in keeping the addresses.

Migration hub · Checklist · Mapping · Tools

1. Read the files, or crawl

You have	Read
The files (FTP, Git) of a hand-made site	the `.html` files directly
A static site generator (Jekyll, Hugo, Eleventy ...)	the **sources** (Markdown with front matter), not the generated HTML; see JavaScript and Node.js apps, 2 for Markdown
Only the live site	a crawl: sitemap, links, server log; save each page

2. Extract the content

1. **Group the pages by template** (pages that share the same layout).
2. **Per template, find the content area** (a `<main>`, an `<article>`, a `<div id="content">`) and the parts to keep: title (`<h1>` or `<title>`), date, author, meta description.
3. **Clean the HTML**: drop navigation, scripts, inline styles and layout tables; keep headings, paragraphs, lists, tables, links and images. Reduce it to Exponential's simplified input (`<p>`, `<strong>`, `<em>`, `<a href>`, `<h1>`-`<h6>`, `<ul>`/`<ol>`/`<li>`, `<table>`/`<tr>`/`<td>`), which the kernel parses into XmlText (tools, section 4).
4. **Images and documents**: collect every `<img src>` and linked PDF/office file, import them as image and file objects, and replace the references with embeds or links to the new objects.
5. **The tree**: the directory structure usually is the navigation: `/products/widgets/index.html` becomes a folder *Products*, a folder *Widgets*, and so on.

3. Import

Write JSON rows (remote id from the old path, for example `html-/products/widgets.html`), then import with the script skeleton or xrowextract's importer. A folder of images and documents can come in through the upload page's ZIP importer, keeping its folder structure.

4. Addresses

Static addresses end in `.html` or are directories with `index.html`. Exponential's aliases are clean paths (`products/widgets`). For each old address, add a redirecting custom URL alias on the new node, and check what the alias cleaning makes of the dot and the extension; when the old addresses follow one pattern, a URL wildcard or a web-server rule (strip `.html`, `index.html`) covers them all (checklist, 4).

5. From Exponential to a static copy

Exponential can write static pages of its public siteaccesses: `./console exp:makestaticcache` crawls them and generates static pages (`--site`, `--max-pages`, `--max-depth`). For an archive copy of a site you are switching off, a crawl with a mirroring tool of the live site works as well; keep the storage directory for the originals of the images.

6. Checklist

- ☐ Address list complete (files, sitemap, log); pages grouped by template.
- ☐ Extractor per template tested; HTML reduced to the simplified input.
- ☐ Images and documents imported; references rewritten.
- ☐ Tree built from the directory structure; remote ids from old paths.
- ☐ `.html` addresses redirected and tested.

> **Become an Exponential Partner**
>
> Partners share their information, provide funding, contribute back and develop the products with us, and that is how Exponential grows. A partner website and dashboard are planned for mid-2027: please become a partner today to speed up our growth rate and land new features faster.
>
> Partners: start here | How to apply

This page is part of the Exponential documentation and is free documentation under the GNU Free Documentation License 1.3 or any later version (the license of this documentation).

Exponential Copyright © 1998-2026 Exponential Foundation and 7x
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".
This documentation is free documentation, licensed under the GNU Free Documentation License 1.3 or later:
https://www.gnu.org/licenses/fdl-1.3.html.