

GitBook

Contents

- Load from single GitBook page
- Load from all paths in a given GitBook

How to pull page data from any GitBook.

```
from langchain.document_loaders import GitbookLoader
```

```
loader = GitbookLoader("https://docs.gitbook.com")
```

Load from single GitBook page

```
page_data = loader.load()
```

```
page_data
```

```
[Document(page_content='Introduction to GitBook\nGitBook is a modern documentation platform where teams can document everything from products to internal knowledge bases and APIs.\nWe want to help \nteams to work more efficiently\n by creating a simple yet powerful platform for them to \nshare their knowledge\n.\nOur mission is to make a \nuser-friendly\n and \ncollaborative\n product for everyone to create, edit and share knowledge through documentation.\nPublish your documentation in 5 easy steps\nImport\n\nMove your existing content to GitBook with ease.\nGit Sync\n\nBenefit from our bi-directional synchronisation with GitHub and GitLab.\nOrganise your content\n\nCreate pages and spaces and organize them into collections\nCollaborate\n\nInvite other users and collaborate asynchronously with ease.\nPublish your docs\n\nShare your documentation with selected users or with everyone.\nNext\n - Getting started\nOverview\nLast modified \n3mo ago', lookup_str='', metadata={'source': 'https://docs.gitbook.com', 'title': 'Introduction to GitBook'}, lookup_index=0)]
```

[Skip to main content](#)

Load from all paths in a given GitBook

For this to work, the GitbookLoader needs to be initialized with the root path

(`https://docs.gitbook.com` in this example) and have `load_all_paths` set to `True`.

```
loader = GitbookLoader("https://docs.gitbook.com", load_all_paths=True)
all_pages_data = loader.load()
```

```
Fetching text from https://docs.gitbook.com/
Fetching text from https://docs.gitbook.com/getting-started/overview
Fetching text from https://docs.gitbook.com/getting-started/import
Fetching text from https://docs.gitbook.com/getting-started/git-sync
Fetching text from https://docs.gitbook.com/getting-started/content-structure
Fetching text from https://docs.gitbook.com/getting-started/collaboration
Fetching text from https://docs.gitbook.com/getting-started/publishing
Fetching text from https://docs.gitbook.com/tour/quick-find
Fetching text from https://docs.gitbook.com/tour/editor
Fetching text from https://docs.gitbook.com/tour/customization
Fetching text from https://docs.gitbook.com/tour/member-management
Fetching text from https://docs.gitbook.com/tour/pdf-export
Fetching text from https://docs.gitbook.com/tour/activity-history
Fetching text from https://docs.gitbook.com/tour/insights
Fetching text from https://docs.gitbook.com/tour/notifications
Fetching text from https://docs.gitbook.com/tour/internationalization
Fetching text from https://docs.gitbook.com/tour/keyboard-shortcuts
Fetching text from https://docs.gitbook.com/tour/seo
Fetching text from https://docs.gitbook.com/advanced-guides/custom-domain
Fetching text from https://docs.gitbook.com/advanced-guides/advanced-sharing-and-security
Fetching text from https://docs.gitbook.com/advanced-guides/integrations
Fetching text from https://docs.gitbook.com/billing-and-admin/account-settings
Fetching text from https://docs.gitbook.com/billing-and-admin/plans
Fetching text from https://docs.gitbook.com/troubleshooting/faqs
Fetching text from https://docs.gitbook.com/troubleshooting/hard-refresh
Fetching text from https://docs.gitbook.com/troubleshooting/report-bugs
Fetching text from https://docs.gitbook.com/troubleshooting/connectivity-issues
Fetching text from https://docs.gitbook.com/troubleshooting/support
```

```
print(f"fetchd {len(all_pages_data)} documents.")
# show second document
all_pages_data[2]
```

[Skip to main content](#)

```
Document(page_content="Import\nFind out how to easily migrate your existing documentation and which formats are supported.\nThe import function allows you to migrate and unify existing documentation in GitBook. You can choose to import single or multiple pages although limits apply. \nPermissions\nAll members with editor permission or above can use the import feature.\nSupported formats\nGitBook supports imports from websites or files that are:\nMarkdown (.md or .markdown)\nHTML (.html)\nMicrosoft Word (.docx).\nWe also support import from:\nConfluence\nNotion\nGitHub Wiki\nQuip\nDropbox Paper\nGoogle Docs\nYou can also upload a ZIP\n\ncontaining HTML or Markdown files when\n\nimporting multiple pages.\nNote: this feature is in beta.\nFeel free to suggest import sources we don't support yet and\n\nlet us know\n\nif you have any issues.\nImport panel\nWhen you create a new space, you'll have the option to import content straight away:\n\nThe new page menu\nImport a page or subpage by selecting\n\nImport Page\n\nfrom the New Page menu, or\n\nImport Subpage\n\nin the page action menu, found in the table of contents:\n\nImport from the page action menu\n\nWhen you choose your input source, instructions will explain how to proceed.\nAlthough GitBook supports importing content from different kinds of sources, the end result might be different from your source due to differences in product features and document format.\nLimits\nGitBook currently has the following limits for imported content:\n\nThe maximum number of pages that can be uploaded in a single import is\n\n20.\n\nThe maximum number of files (images etc.) that can be uploaded in a single import is\n\n20.\n\nGetting started -\n\nPrevious\n\nOverview\n\nNext\n\n- Getting started\n\nGit Sync\n\nLast modified\n\n4mo ago", lookup_str='', metadata={'source': 'https://docs.gitbook.com/getting-started/import', 'title': 'Import'}, lookup_index=0)
```