

[to PDF](#) ▶

# PowerPoint

## Contents

- Retain Elements

This covers how to load PowerPoint documents into a document format that we can use downstream.

```
from langchain.document_loaders import UnstructuredPowerPointLoader
```

```
loader = UnstructuredPowerPointLoader("example_data/fake-power-point.pptx")
```

```
data = loader.load()
```

```
data
```

```
[Document(page_content='Adding a Bullet Slide\n\nFind the bullet slide\nlayout\n\nUse _TextFrame.text for first bullet\n\nUse _TextFrame.add_paragraph()\nfor subsequent bullets\n\nHere is a lot of text!\n\nHere is some text in a text\nbox!', lookup_str='', metadata={'source': 'example_data/fake-power-point.pptx'},\nlookup_index=0)]
```

## Retain Elements

Under the hood, Unstructured creates different “elements” for different chunks of text. By default we combine those together, but you can easily keep that separation by specifying

```
mode="elements"
```

```
loader = UnstructuredPowerPointLoader("example_data/fake-power-point.pptx",
```

[Skip to main content](#)

```
mode="elements")
```

```
data = loader.load()
```

```
data[0]
```

```
Document(page_content='Adding a Bullet Slide', lookup_str='', metadata={'source':  
'example_data/fake-power-point.pptx'}, lookup_index=0)
```