Data is held in different ways – sites can vary from simple to highly complex:

- **Tabular**: Well-organized and easy to extract from [a single page](#) or [100s of pages](#)

- **Vast**: Well-organized but with data scattered [across hundreds of pages.](#)

- **Downloads:** Visit page and [download files](#).

- **Static Links:** Should be easy, but where's the page content?

- **Scrollers/Loan More:** More and more data appears [as you scroll on single page](#).

- **Inputs or Interactivity:** Data is returned only [after some interaction](#).

- **Asynchronous**: Well-organized but reveals [data on request](#).

- **APIs:** Application Programming Interface which allow us to talk directly to the data.

- **Locked-down**: None of the techniques above work. It's robot time!