

# Re: fast method accessing large, simple structured data

---

*Source:* <http://coding.derkeiler.com/Archive/Python/comp.lang.python/2008-02/msg00245.html>

---

- *From:* "Diez B. Roggisch" <[deets@xxxxxxxxxxxxxx](mailto:deets@xxxxxxxxxxxxxx)>
  - *Date:* Sat, 02 Feb 2008 21:51:46 +0100
- 

agc schrieb:

Hi,

I'm looking for a fast way of accessing some simple (structured) data.

The data is like this:

Approx 6 – 10 GB simple XML files with the only elements I really care about are the <title> and <article> ones.

So what I'm hoping to do is put this data in a format so that I can access it as fast as possible for a given request (http request, Python web server) that specifies just the title, and I return the article content.

Is there some good format that is optimized for search for just 1 attribute (title) and then returning the corresponding article?

I've thought about putting this data in a SQLite database because from what I know SQLite has very fast reads (no network latency, etc) but not as fast writes, which is fine because I probably won't be doing much writing (I won't ever care about the speed of any writes).

So is a database the way to go, or is there some other, more specialized format that would be better?

Database it is. Make sure you have proper indexing.

Diez

.