

Ten. Million. Questions. Let's celebrate [all we've done together](#).

Stack Overflow is a question and answer site for professional and enthusiast programmers. It's 100% free.

Take the 2-minute tour

## Ready-made database for Neo4j

asked 4 months ago

viewed 84 times

active 3 months ago



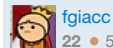
I need to run some tests on Neo4j performances with different sized databases. I would like to know if there are some ready-made databases / Cypher CREATE queries I could use. I tried a workaround: converting an XML database into a CREATE query with the XML to Cypher Converter by Nigel Small but I've not been able to make it work (500 Internal Server Error).

Any suggestions?

xml database neo4j cypher

share improve this question

asked Apr 7 at 11:11



I've found Neo4j website hosts some example dataset to use [here](#). The movie database works like a charm. I also downloaded the Musicbrainz database but it seems there must be something wrong with the archive: every time I try to extract it some errors are thrown (corrupted file or unsupported compression). I also tried to repair it with no success. A database with millions of nodes is exactly what I was looking for! Any idea how to solve this? Any other place I could find it? – [fgiacc](#) Apr 17 at 10:36

add a comment

### 3 Answers

active oldest votes



I don't know of ready-made XML databases, but you can use the utility [xml2neo](#) to transform any XML document you want into a series of cypher statements prepped for neo4j.

At that point, you can then just import whatever XML you want into a neo4j database, and continue until you get the size that you want. For example, you could easily pull any of millions of ATOM or RSS feeds that way.

share improve this answer

answered Apr 7 at 12:47



This tool could be great but it seems to have some flaws. I tried to use it with the sample.xml provided and it generated some cypher queries that threw syntax errors when used in Neo4j. – [fgiacc](#) Apr 8 at 8:46

Post it as an issue along with a sample of your xml and it will get corrected – [FrobberOfBits](#) Apr 8 at 15:05

add a comment

#### Blog

[Why Stack Overflow is a Good Workplace for Women](#)

#### Related

- 1 [Import a Simple Graph into Neo4j](#)
- 0 [Neo4J 2.0 How to refactor the database to introduce Date nodes](#)
- 0 [Neo4j Cypher query fails and return with an Unknown Error](#)
- 2 [Neo4j cypher for co-favorited plus additional attribute](#)
- 1 [Neo4j create multiple nodes and relationships](#)
- 0 [Mutual Friends of Friends Neo4j Cypher](#)
- 0 [Deploying Neo4j database](#)
- 1 [Neo4j Huge database query performance configuration](#)
- 2 [Make my Neo4j queries faster](#)
- 0 [Howto recover from neo4j cypher error: Neo.ClientError.Statement.ENTITYNotFound](#)

#### Hot Network Questions

- [Has any Star Trek character appeared in all three division uniform colors?](#)
- [Why did Holly decide to simply wait out the radioactivity?](#)
- [Why don't decoupling/bypass capacitors need resistors to perform their function, like regular filters?](#)
- [Do companies only pay dividends if they are in profit?](#)
- [Could a communication satellite that fails to reach GEO be repurposed?](#)
- [How to use Piano pedals without using the foot?](#)
- [I added more water to my ciabatta dough by accident](#)
- [Suggested alternatives for that horrible new noun 'nice-to-have'?](#)



It's not "ready-made" persay where it installs out of the box with a fully populated database, however I am getting more interested in Neo4j's spatial capabilities and am using the following as my base to start from.

<http://www.lyonwj.com/mapping-the-worlds-airports-with-neo4j-spatial-and-openflights-part-1/>

In this site, it provides the link to the dataset of worldwide airports (though it's a bit difficult to find at first, so the specific link to the dataset is below):



<https://sourceforge.net/p/openflights/code/HEAD/tree/openflights/data/airports.dat?format=raw>

The blog also provides the Python script to perform the insertion, though you could easily insert via a LOAD CSV Cypher script (I suggest adding headers that match the insertion below first; e.g. airportId, airportName, airportCity, etc.):

```
USING PERIODIC COMMIT 500
LOAD CSV WITH HEADERS FROM "file:d:/airports.dat" AS line
CREATE (airport:Airport { id: line.airportId, name: line.airportName, city: line.air
```

More information about the contents of the dataset is available here:

<http://openflights.org/data.html>

The blog posting above also then provides details on querying via Cypher, Java, etc.

If I come across any others, I'll update my answer.

**EDIT 1** The dataset is not terribly big (less than 10,000 nodes), so you likely don't need the USING PERIODIC COMMIT 5000 or you can lower that number significantly if you wanted to keep it.

Also, this dataset is just nodes, and doesn't have any relationships.

You could certainly create relationships though, such as instead of inserting the country name into the airport node, you could create a country node and then relate the airport nodes to their associated country node:

Don't quote me on this (it may well be syntactically incorrect and/or not the best way to do this in Neo4j, but should give you a start) as I am still very new to Cypher, but I think it would go something like:

```
USING PERIODIC COMMIT 500
LOAD CSV WITH HEADERS FROM "file:d:/airports.dat" AS line
CREATE (airport:Airport { id: line.airportId, name: line.airportName, city: line.air
WITH airport, MERGE (country:Country { name: line.airportCountry })
ON CREATE SET (airport)-[:IN_COUNTRY]->(country)
ON MATCH SET (airport)-[:IN_COUNTRY]->(country);
```

**EDIT 2** Additionally, check out:

<http://maxdemarzi.com/2014/01/31/neo4j-spatial-part-1/>

Specifically the reference to the restaurant data you can implement.

share improve this answer

edited Apr 7 at 17:33

answered Apr 7 at 14:37



Brooks

113 • 11

The airports dataset doesn't fit my needs. Though I could add some relationships as you suggested, this would create a bunch of clusters instead of one, big, connected graph that could instead be used for path-finding performance tests. Anyway Neo4j Spatial is an interesting project I didn't know and I'll keep an eye on it => - [fgiacc](#) Apr 8 at 9:04

add a comment



0



I've managed to find a way to import my .xml into Neo4j. I converted the .xml into a .csv and then I imported it following [this instructions](#).

share improve this answer

answered Apr 24 at 16:50



fgiacc

22 • 5

add a comment

Your Answer



Is it 'Batman' or 'Batmans'?



Does developing on ARM Processors from different vendors differ much?



Preventing a Print expression from generating a new cell each time it executed



What can I do when I accidentally gave out an overpowered item?



Does Lebanon ban the diary of Anne Frank?



Does adult neurogenesis occur only in the olfactory bulb and the hippocampus?



How can I grant a non-Super User group permission to force check-in of articles?



Legacy application unreachable code



Where did the nickname of 'Bones' for McCoy come from?



Sign that word!



Draw horizontal lines like open and close brackets



Can a plane fly backward if its prop's in reverse?



What does OWA stand for?



Should you always minimize cognitive load






When does the designer pay for a reprint



Did a Kamikaze ever impact the hull of a ship, as opposed to the deck?

Sign up or [log in](#)

Post as a guest

-  Sign up using Google
-  Sign up using Facebook
-  Sign up using Stack Exchange

<b>Name</b>	<input type="text"/>
<b>Email</b>	<input type="text"/>

Post Your Answer

By posting your answer, you agree to the [privacy policy](#) and [terms of service](#).

Not the answer you're looking for? Browse other questions tagged `xml` `database` `neo4j` `cypher` or [ask your own question](#).



TECHNOLOGY			LIFE / ARTS	CULTURE / RECREATION	SCIENCE	OTHER
Stack Overflow	Programmers	Database Administrators	Photography	English Language & Usage	Mathematics	Stack Apps
Server Fault	Unix & Linux	Drupal Answers	Science Fiction & Fantasy	Skeptics	Cross Validated (stats)	Meta Stack Exchange
Super User	Ask Different (Apple)	SharePoint	Graphic Design	Mi Yodeya (Judaism)	Theoretical Computer Science	Area 51
Web Applications	WordPress Development	User Experience	Movies & TV	Travel	Physics	Stack Overflow Careers
Ask Ubuntu	Geographic Information Systems	Mathematica	Seasoned Advice (cooking)	Christianity	MathOverflow	
Webmasters		Salesforce		Arqade (gaming)	Chemistry	
Game Development	Electrical Engineering	ExpressionEngine®	Home Improvement	Bicycles	Biology	
TeX - LaTeX	Android Enthusiasts	Answers	Personal Finance & Money	Role-playing Games	more (5)	
	Information Security	more (13)	Academia	more (21)		
			more (9)			