



XML and RDF Architectural Beastiary
Integration Patterns for Documents and Graphs

Charles Greer
XML Prague 2013

The Point

Large enterprises have determined that XML databases are excellent for large scale search and also for full-stack applications.

But for managing disparate, complex business data, they're turning to RDF and triple stores.

In what ways can these technologies complement each other in a large IT shop?



Context and Me

- Enterprise IT management
- Databases (XML databases and triple stores)
- RESTful architectures

- Slavistiki (Academia)
- Geographic Information Systems (government)
- Software Development and IT management (Publishing)
- Software Engineering (Database Software)



Query, Response, RDF.

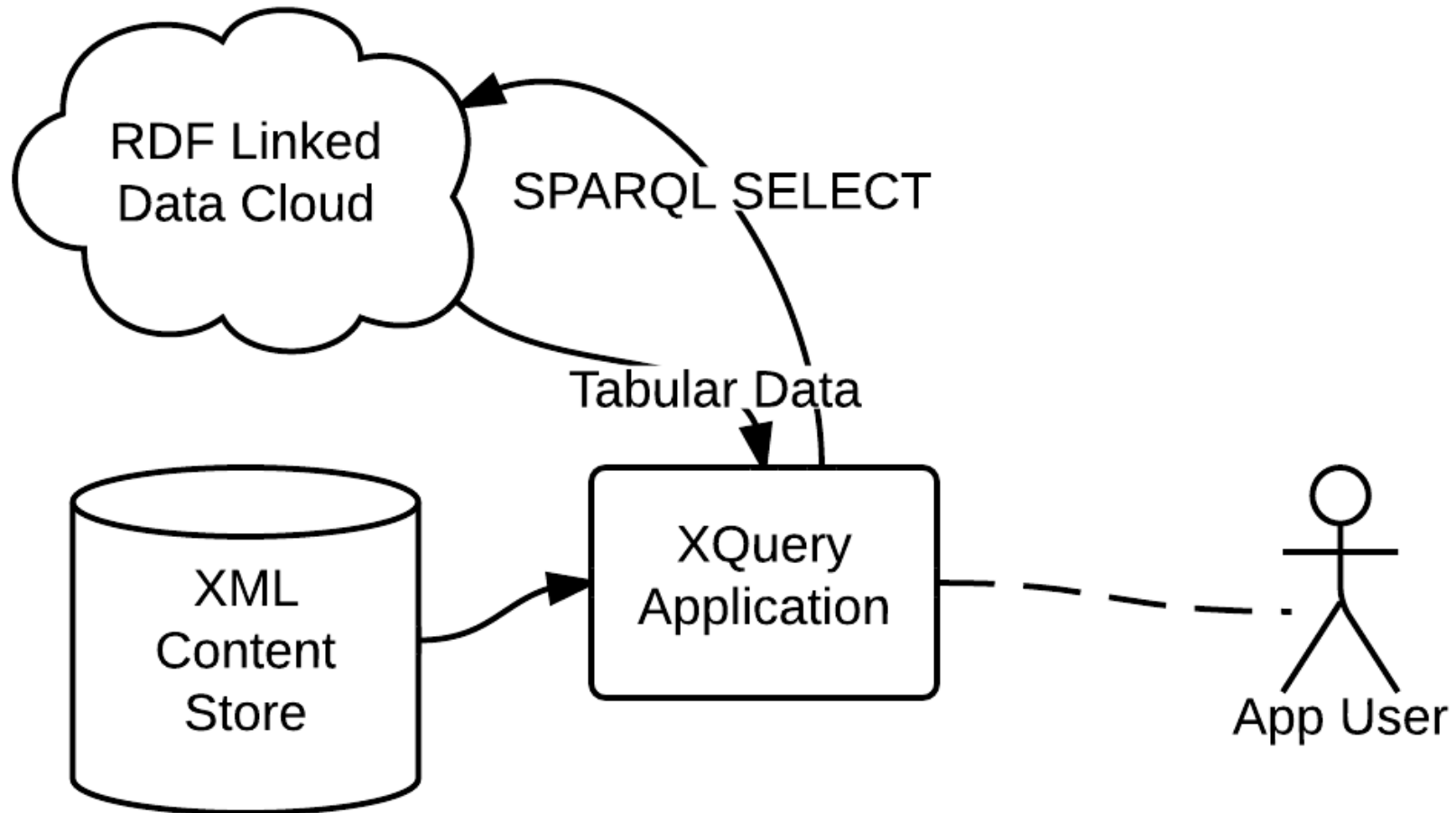
```
SELECT ?s ?p ?o {?s ?p ?o. filter (?s = <http://BK>)}
```

```
<response xmlns="http://www.w3.org/Protocols/rfc2616">  
... <binding name="s"><uri>http://BK</uri></binding>  
    <binding name="p"><uri>dc:title (sic)</uri></binding>  
    <binding name="o">The Brothers Karamazov</binding>...  
</response>
```

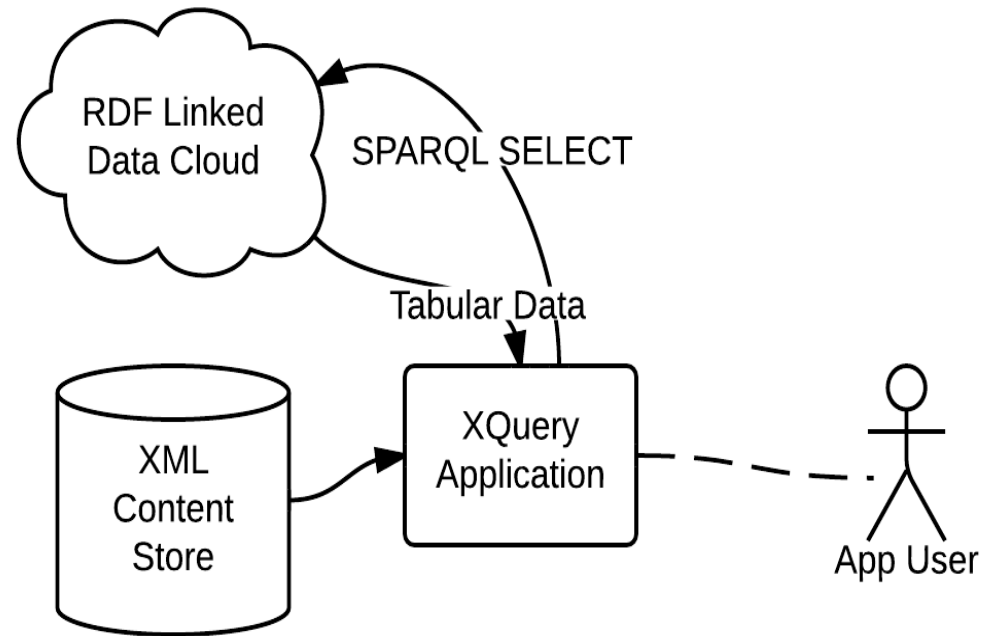
```
<rdf:Description rdf:about="http://BK">  
    <dc:title>The Brothers Karamazov</dc:title>  
</rdf:Description>
```



Beast The First: The Consumer

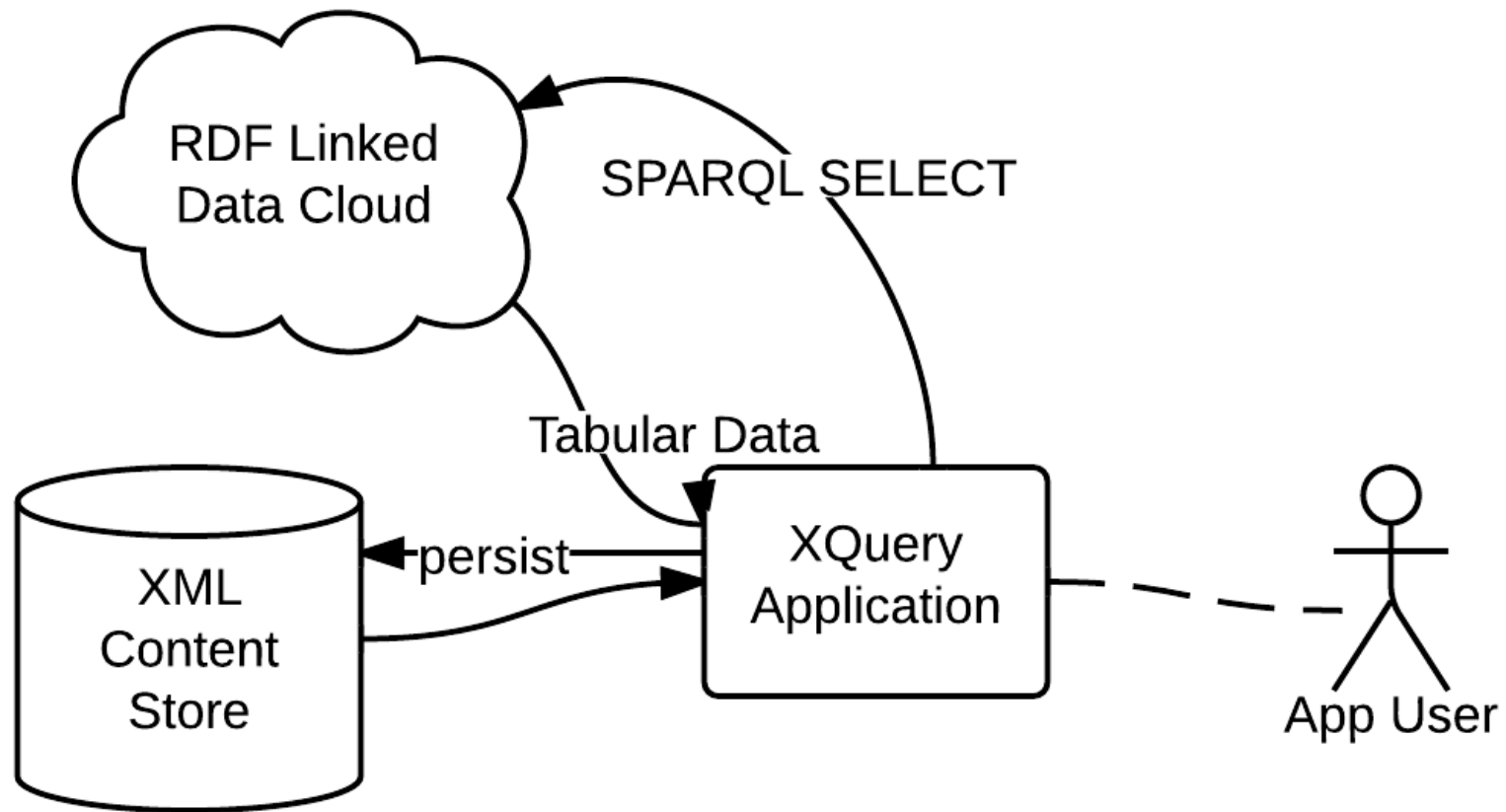


Beast The First: The Consumer

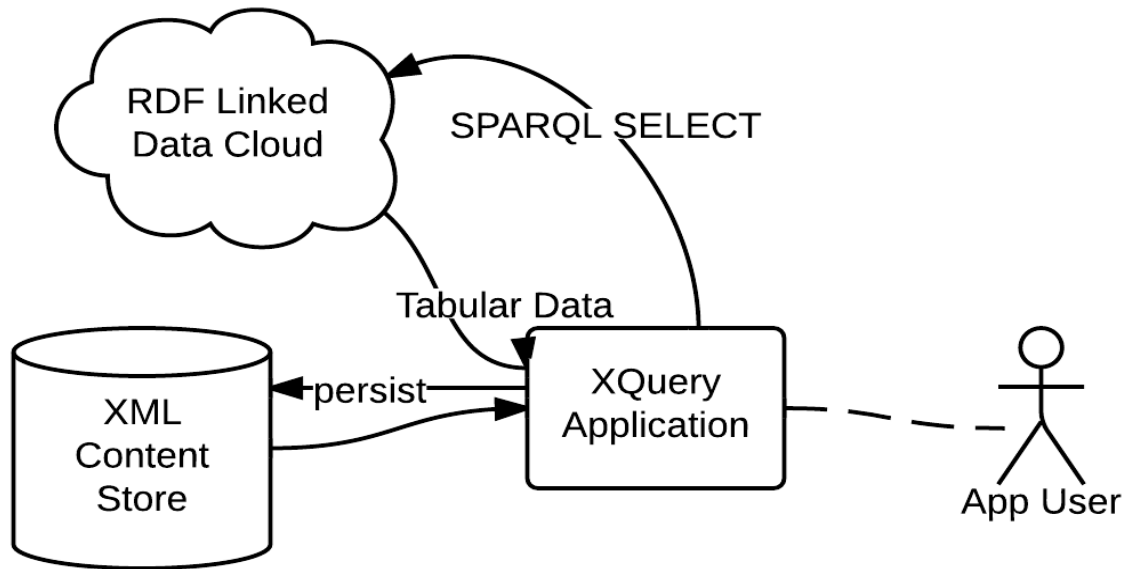


- Uses available systems
- Incremental requirement for data integration?

Beast the Second: The Ingester

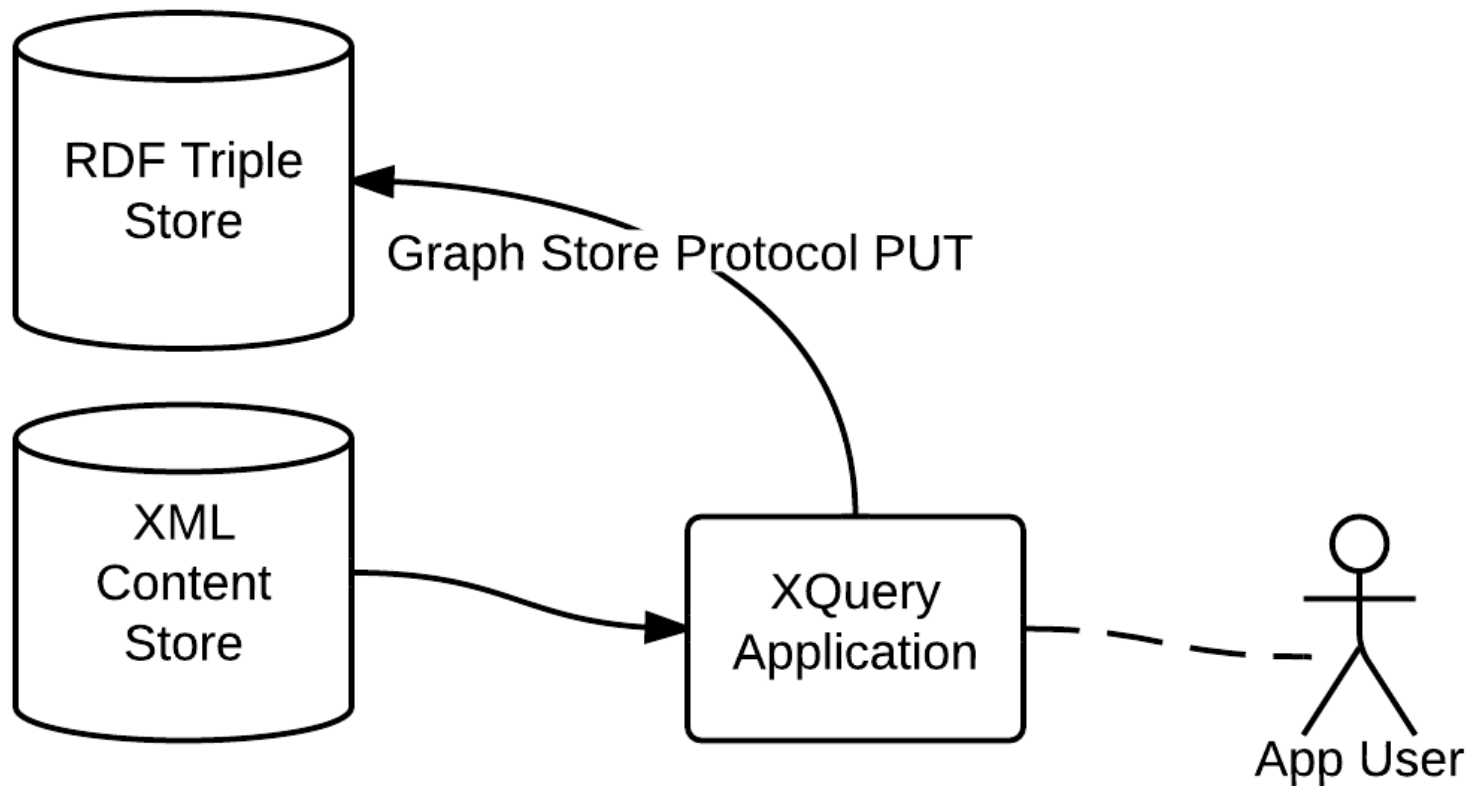


Beast the Second: The Ingester

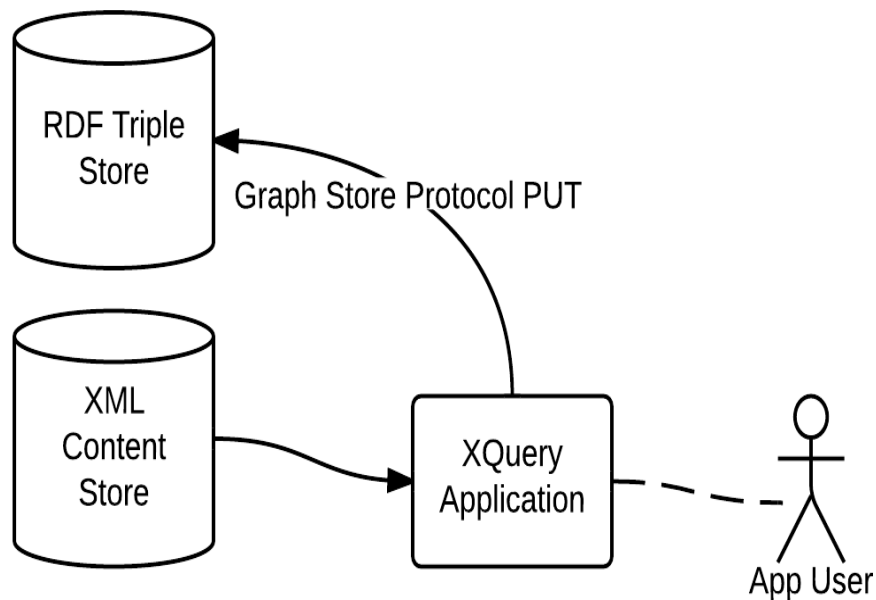


Integration of new domain objects

Beast the Third: The Publisher

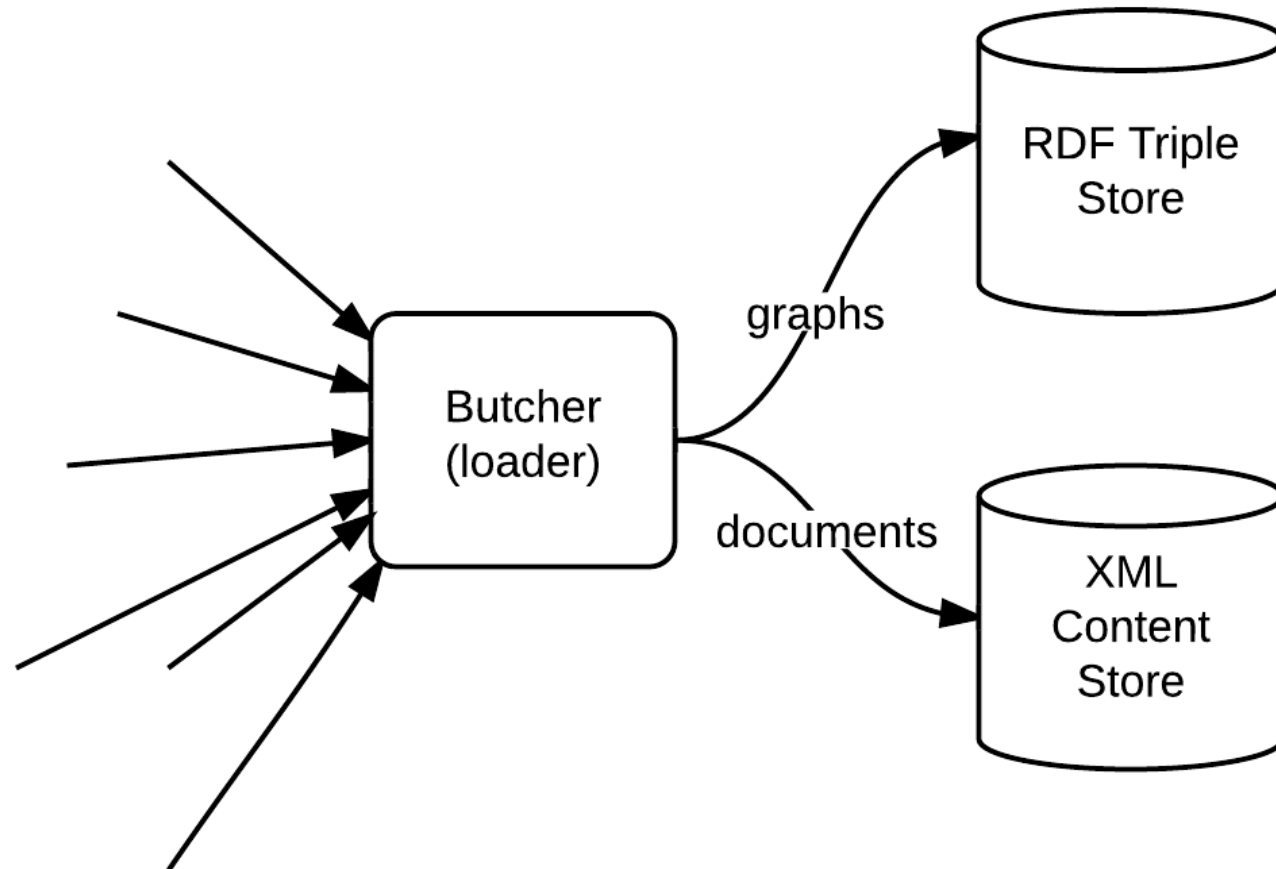


Beast the Third: The Publisher

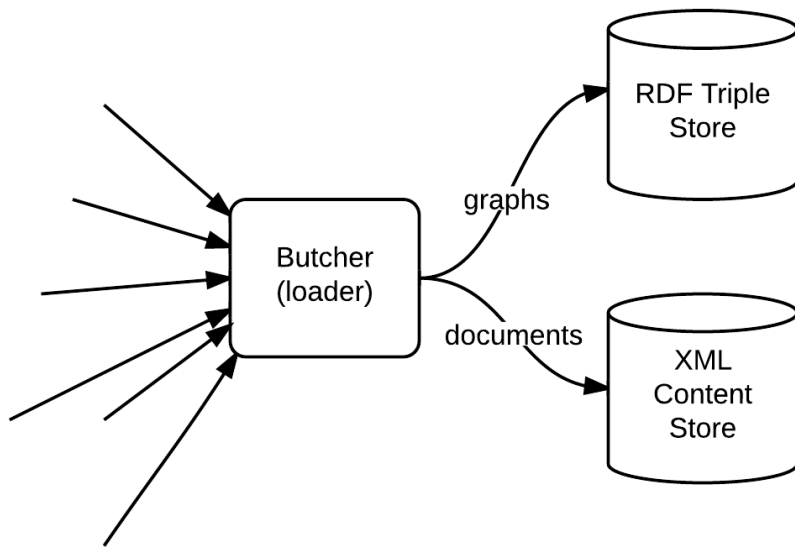


- Requires transforms to RDF
- Meets requirement for data exports from XML database
- Meets requirement for publishing standard data

Beast the Fourth: The Butcher



Beast the Fourth: The Butcher



- Manages updates of XML and RDF data
- Maintains domain objects with little temporal skew
- Provides single source for orchestrated, distributed resources.

Beast the Fifth: The URI

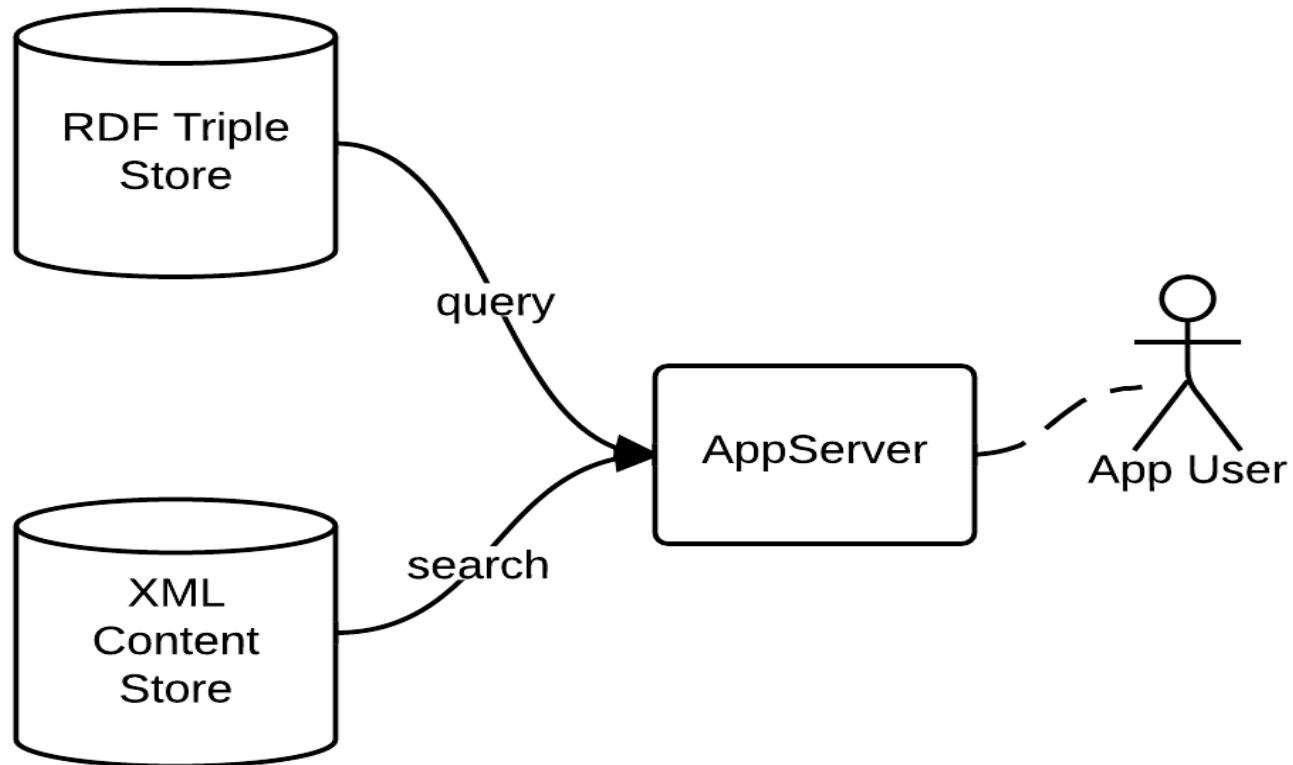
- `<http://the/most/important/takeaway/is/the/uri>`
- `<http://the/most/important/takeaway/is/the/uri>`
- `<http://the/most/important/takeaway/is/the/uri>`
- `<http://the/most/important/takeaway/is/the/uri>`
- `<http://the/most/important/takeaway/is/the/uri>`
- `<http://the/most/important/takeaway/is/the/uri>`

Beast the Fifth: The URI

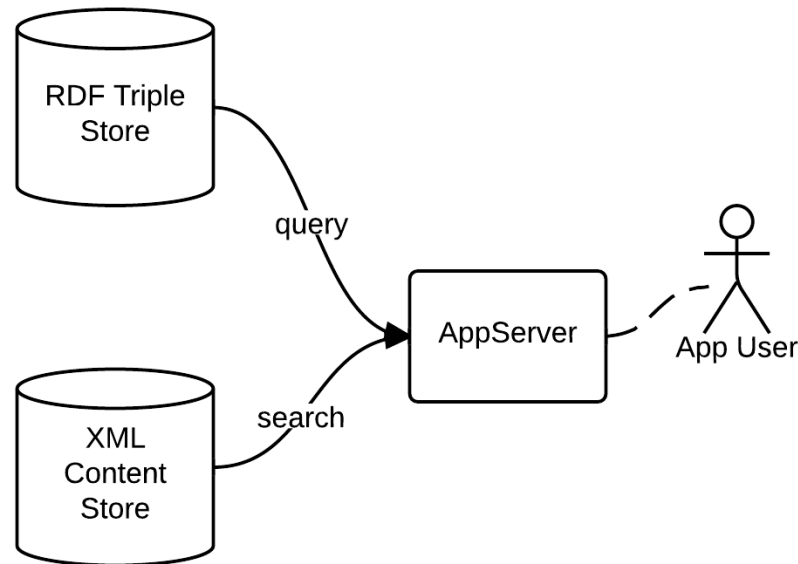
- Identifies data universally
- Atom of RDF and Linked Data worlds
- Convenient alignment of web and data standards
- Goes hand-in-hand with Open World Assumption
- Basis for XML and RDF data integration scenarios

- Has awkward relationship with QName

Beast the Sixth (and Seventh): N-tier Apps



Beast the Sixth (and Seventh): N-tier Apps



- RDF object model with document search
- XML model with data integrated from other business units.
- Generalized query and search services

Beast the Sixth (and Seventh): N-tier Apps

Ebooks | **Print Books** | **Videos** | **DVDs** | **Other**

Sort by: [Date](#) | [Alphabetical](#)

You get lifetime access to ebooks you purchase through oreilly.com. Whenever possible we provide them to you in five DRM-free file formats — PDF, ePub, Kindle-compatible .mobi, DAISY, and Android .apk — that you can use on the devices of your choice. Our ebooks are enhanced with color images, even when the print version is black and white. They are fully searchable, and you can cut-and-paste and print them. We also alert you when we've updated your ebooks with corrections and additions. Now includes **Dropbox Sync!**

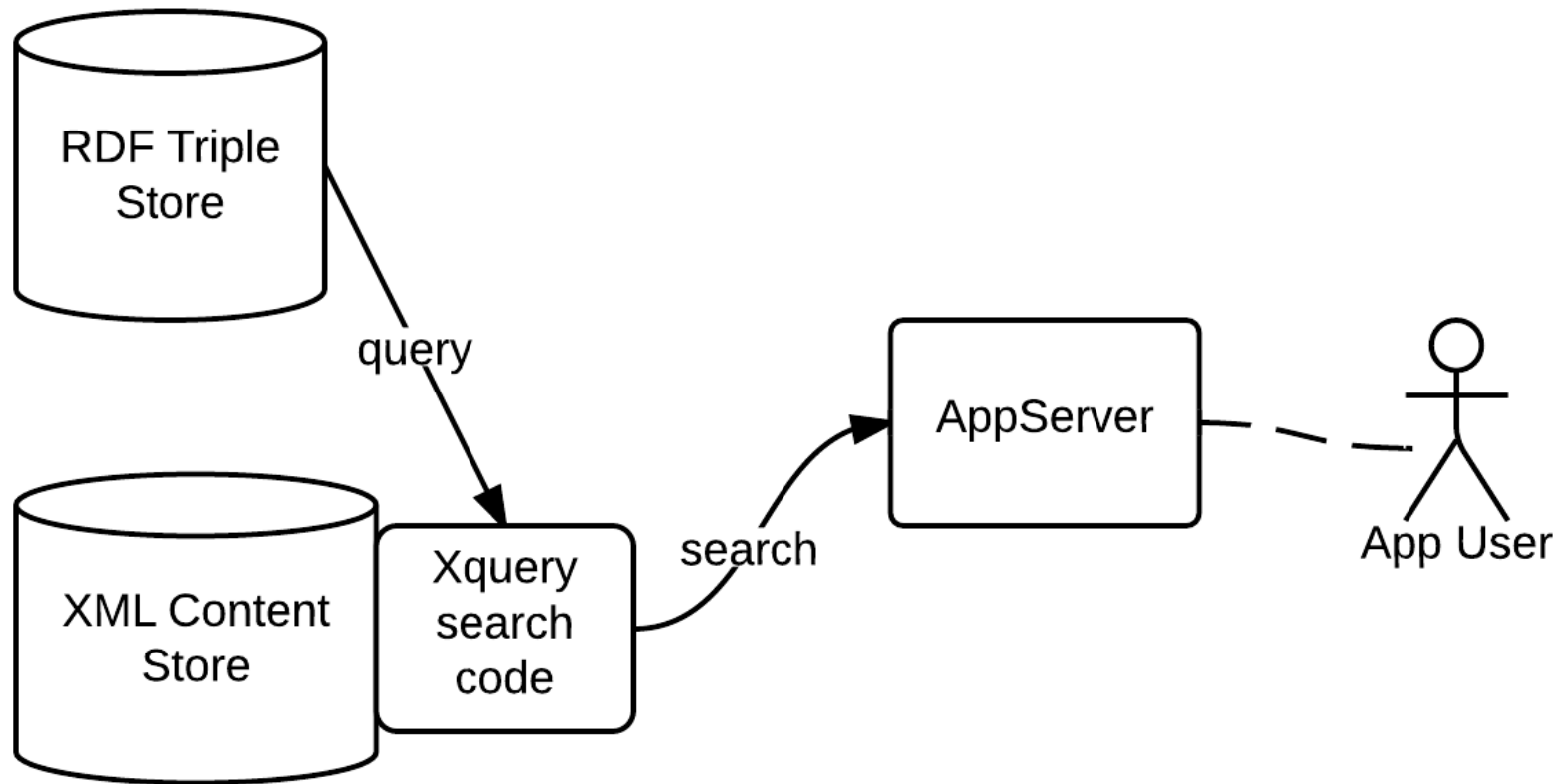
Please contact customer service if you have any questions. 1-800-998-9938 / 707-827-7000 / accounts@oreilly.com

The Art of Application Performance Testing 40% off print edition
By Ian Molyneux
Purchased / Registered: June 2011
[DAISY](#) [ePub](#) [Mobi](#) [PDF](#)
[Errata](#)

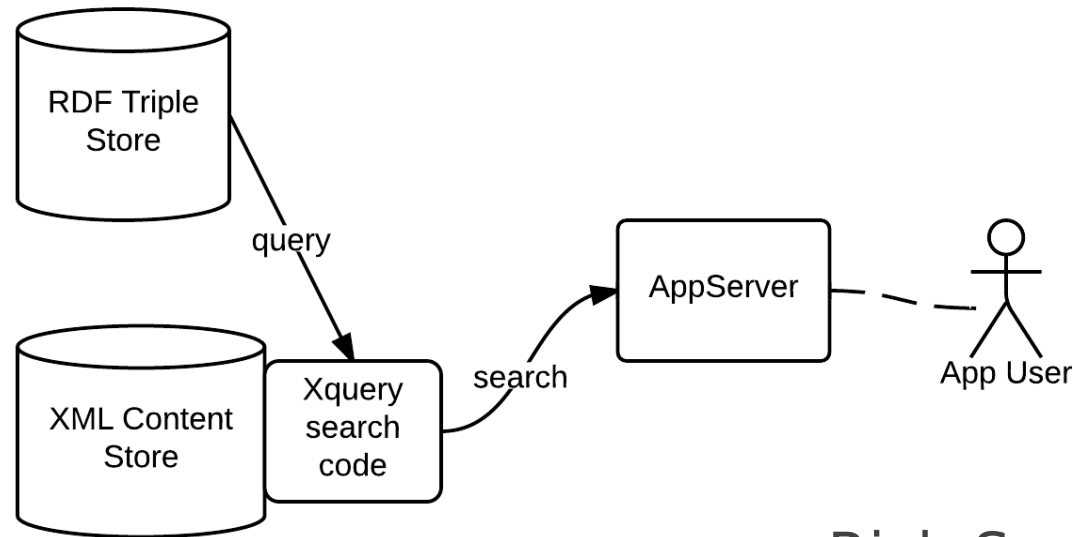
The Book of Audacity
By Carla Schroder
Purchased / Registered: March 2011
[ePub](#) [Mobi](#) [PDF](#)

Data Analysis with Open Source Tools 40% off print edition
By Philipp K. Janert
Purchased / Registered: October 2010
NEW [APK](#) [DAISY](#) [ePub](#) [Mobi](#) [PDF](#)
[Errata](#)

Beast the Eighth: Toward a Monolith



Beast the Eighth: Toward a Monolith



- Rich Search service with embedded queries, output transforms.

Beast the Ninth: Single Repository



Beast the Tenth: RDF within XML

<t>

<s>http://dbpedia.org/resource/The_Brothers_Karamazov</s>

<p><http://www.w3.org/2000/01/rdf-schema#label></p>

<o xml:lang="en">The Brothers Karamazov</o>

</t>

Beast the Tenth: RDF within XML

```
<book
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/terms/">
  <rdf:Description rdf:about="http://BK">
    <dc:title>The Brothers Karamazov</dc:title>
  </rdf:Description>
</part><section></section></part></book>
```

Beast the Tenth: RDF within XML

```
<body vocab="http://purl.org/dc/terms/">
```

```
...
```

```
<div resource="http://BK">
```

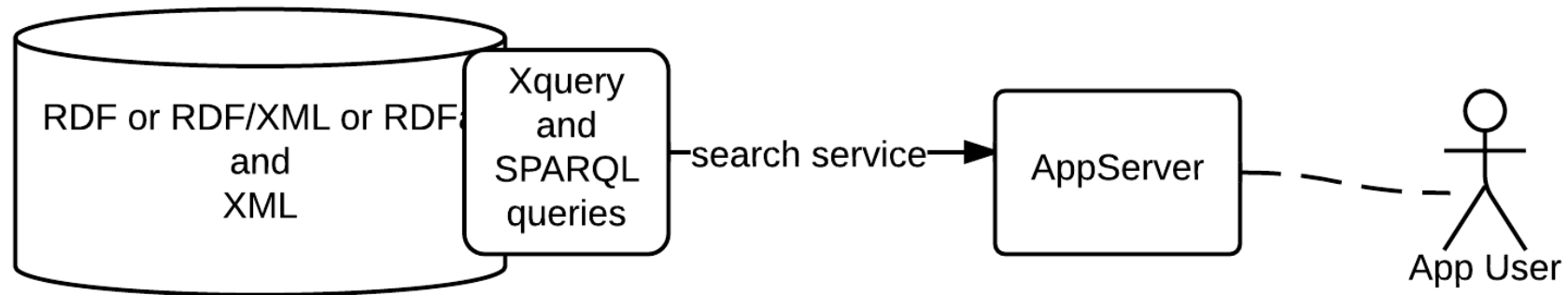
```
<h2 property="title">The Brothers Karamazov</h2>
```

```
</div>
```

```
...
```

```
</body>
```

Beast the Eleventh: The Gleaner



XML practitioners!

You will lead businesses into
a new era of content and
data management.



Thank You!

Charles Greer
cgreer@marklogic.com
@grechaw