

# **XML Databases**

## **Who cares?**



Mary Fernández  
AT&T Labs - Research

# XML Data-Exchange Illusion



***XML  
Documents  
& Schema***

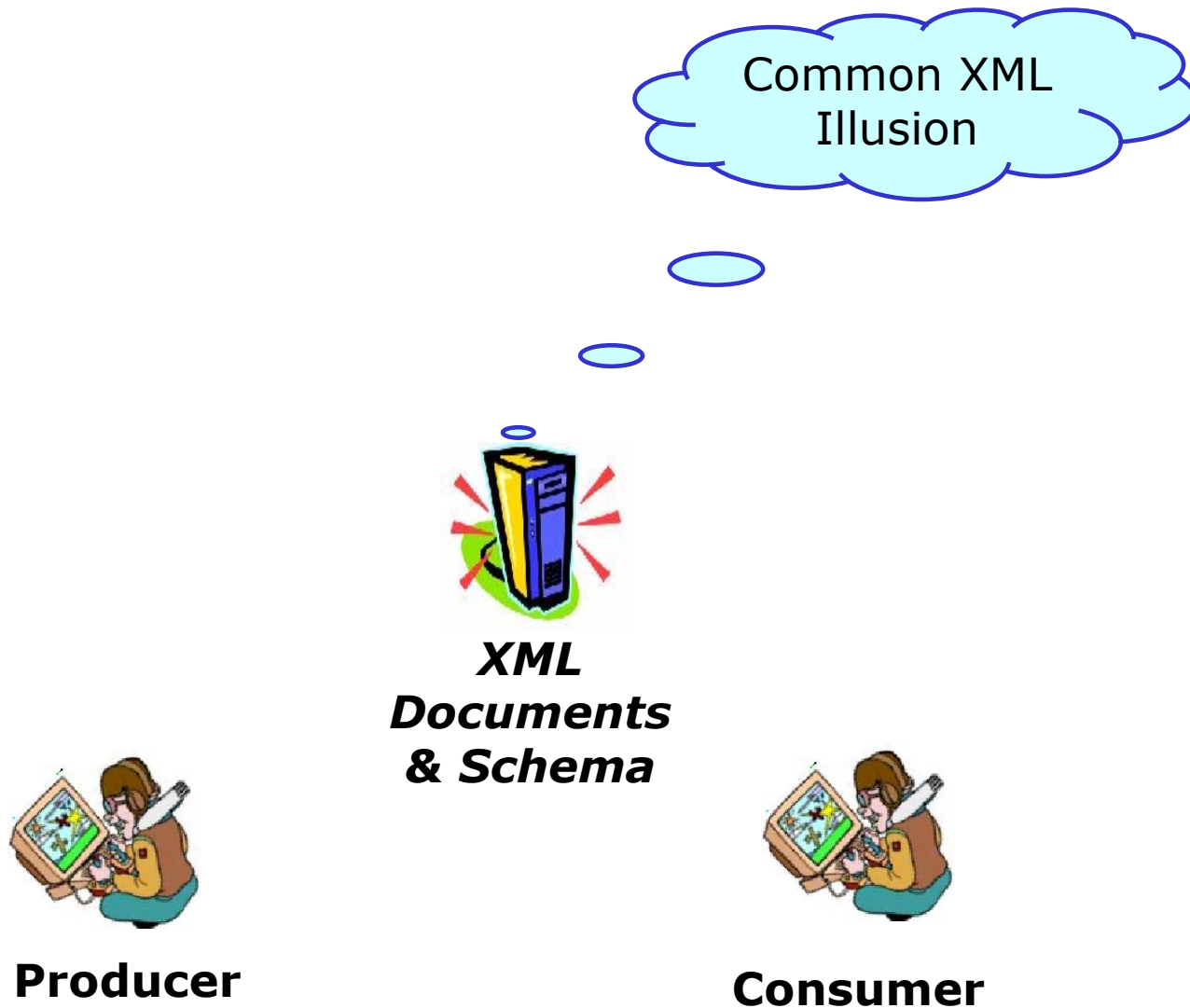


**Producer**

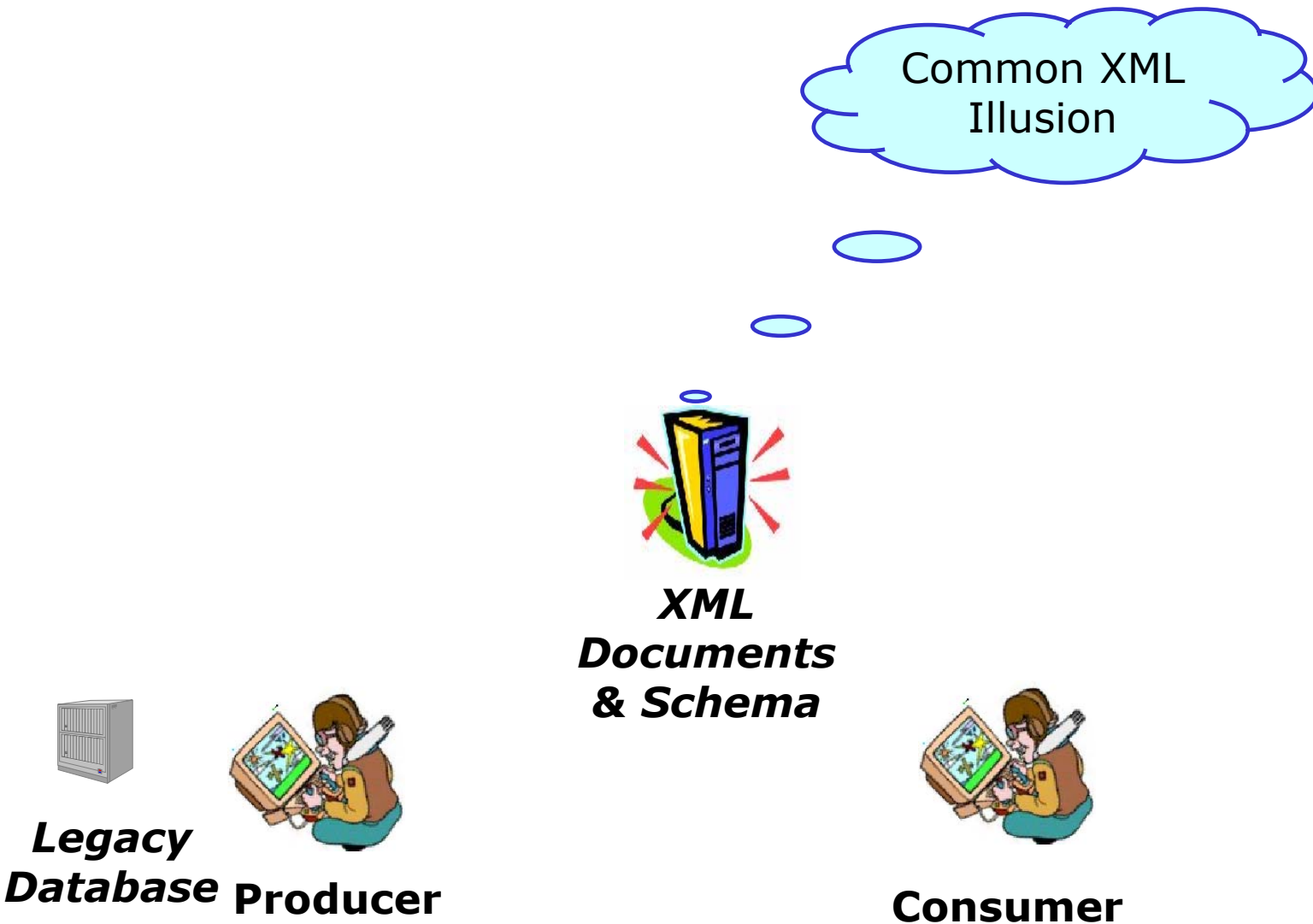


**Consumer**

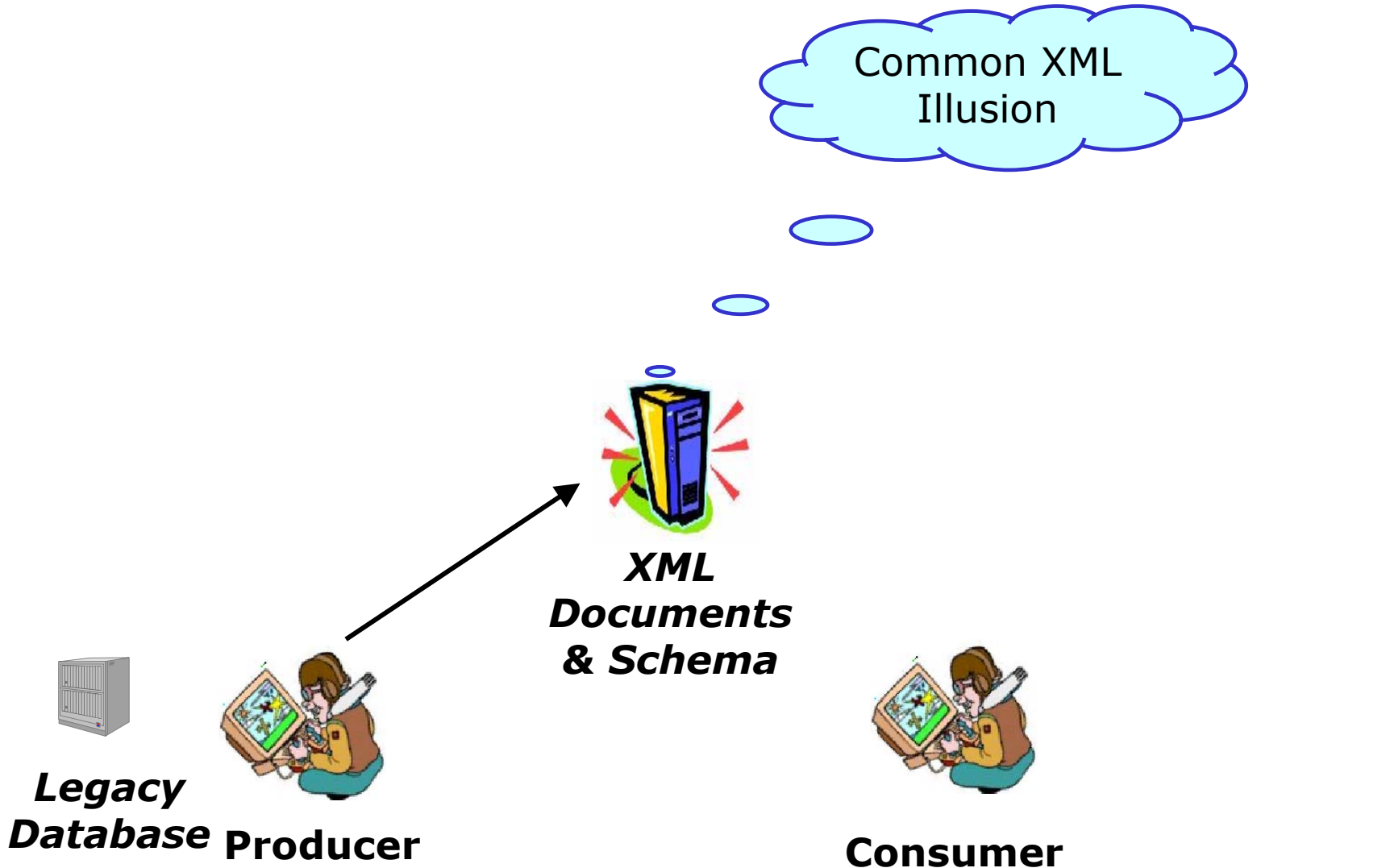
# XML Data-Exchange Illusion



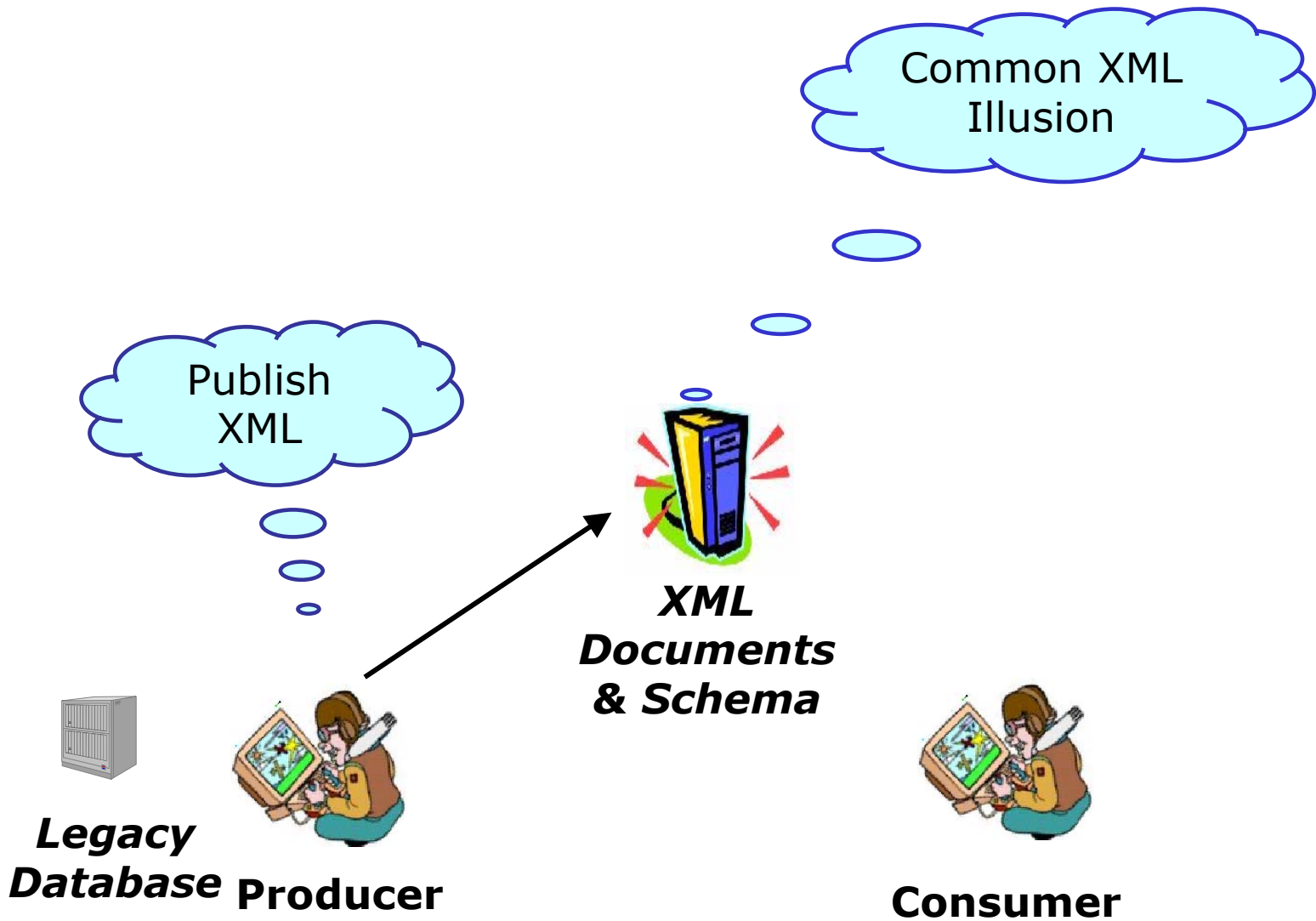
# XML Data-Exchange Illusion



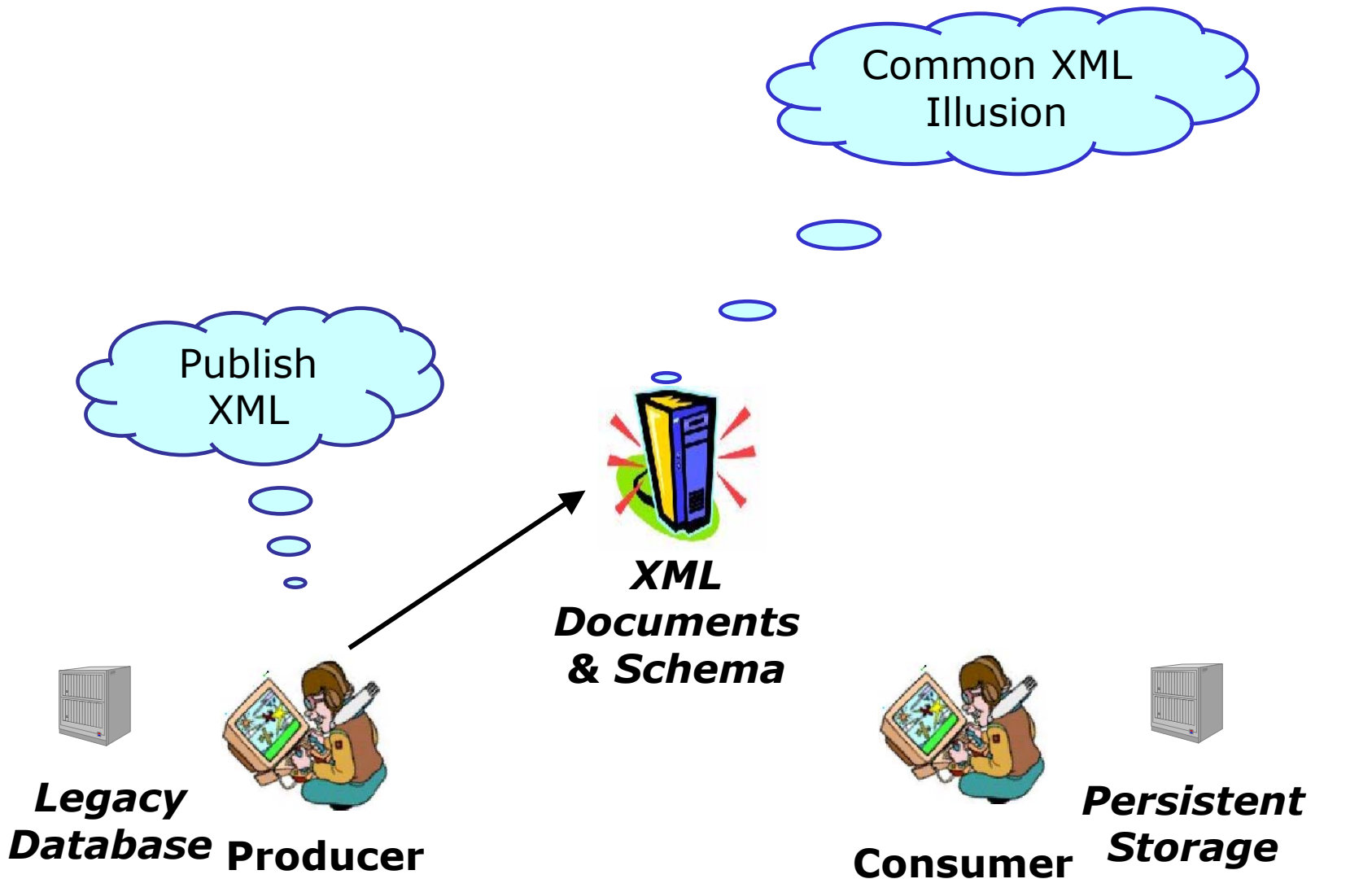
# XML Data-Exchange Illusion



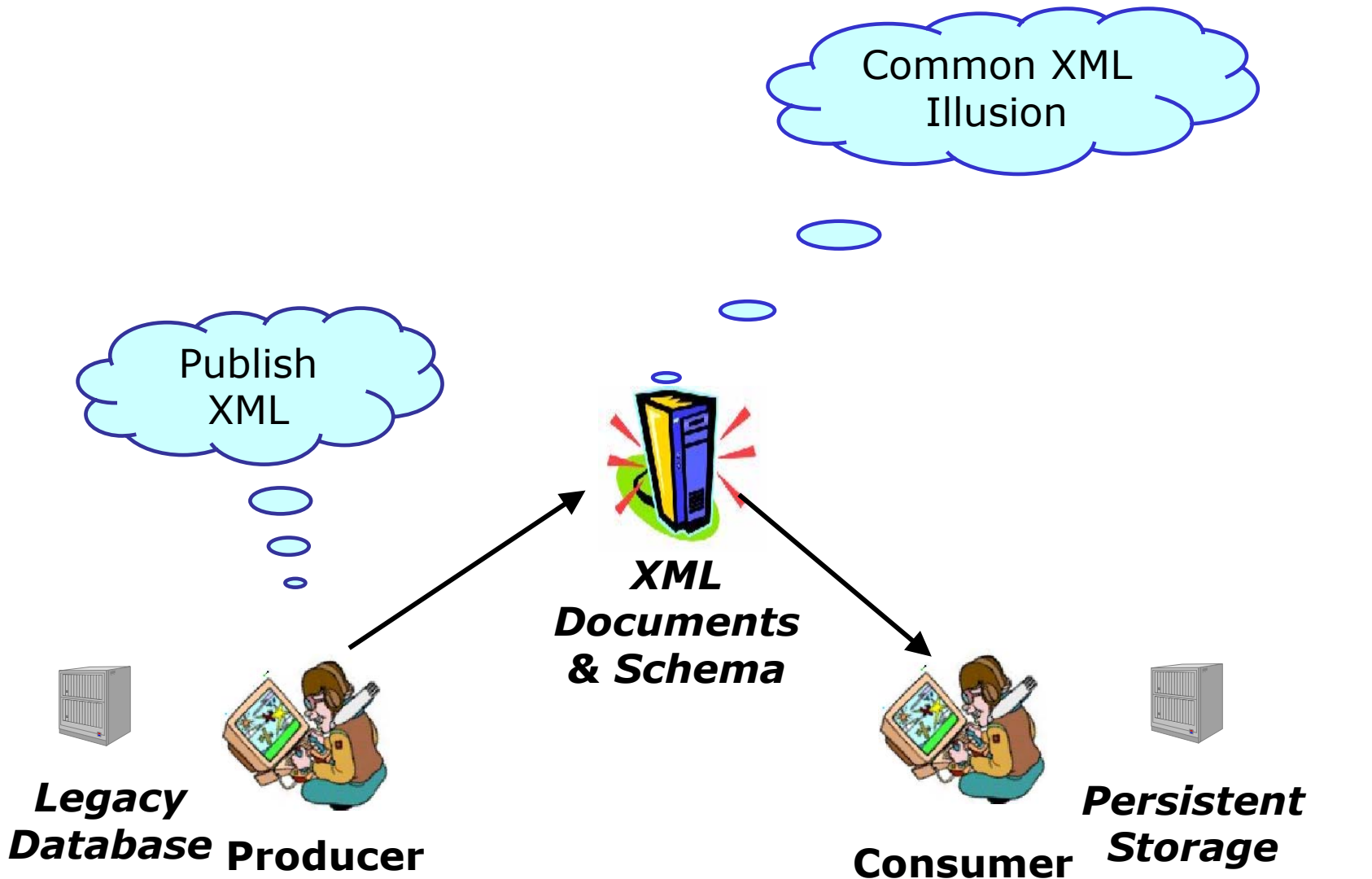
# XML Data-Exchange Illusion



# XML Data-Exchange Illusion

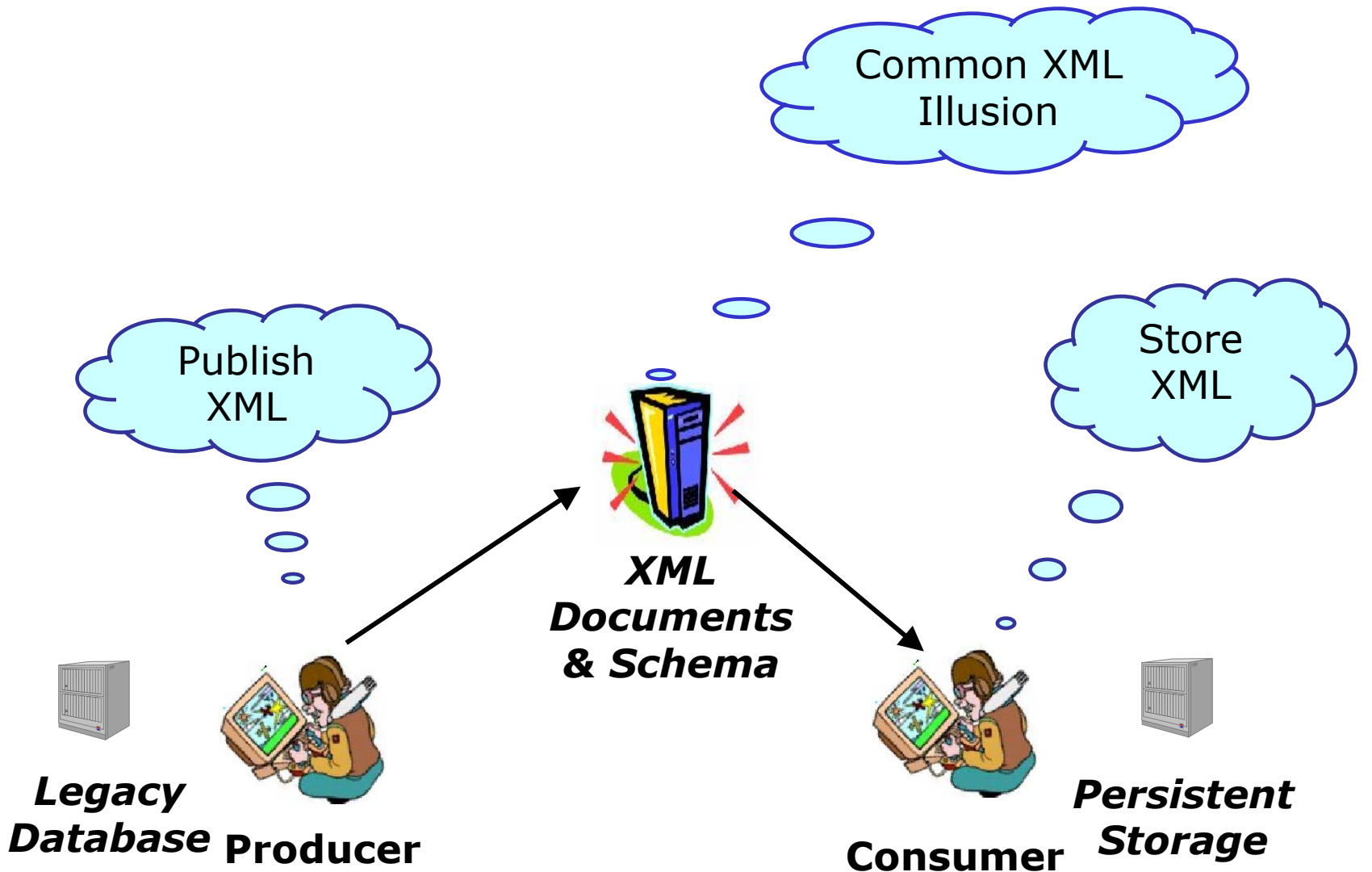


# XML Data-Exchange Illusion

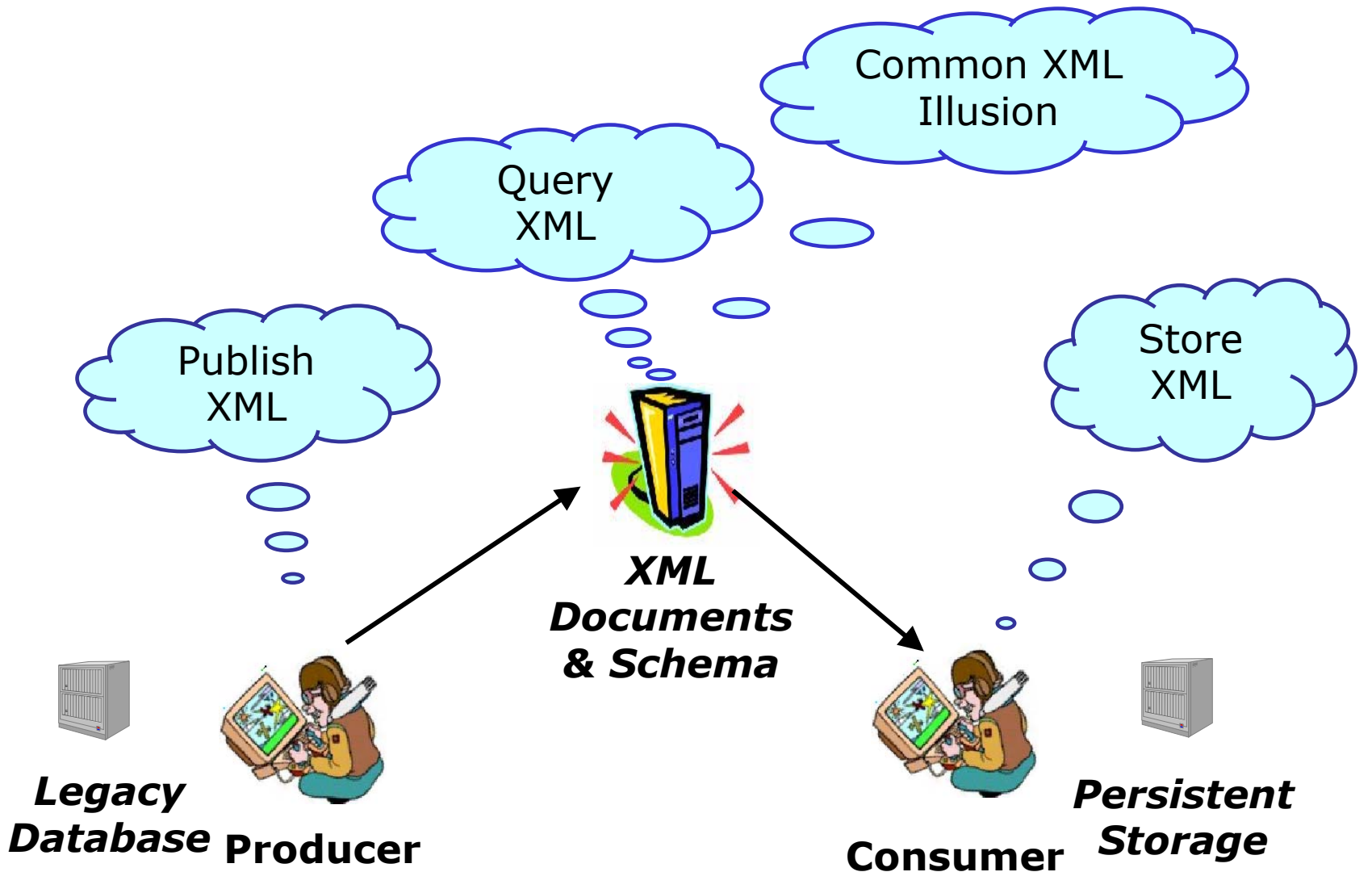




# XML Data-Exchange Illusion



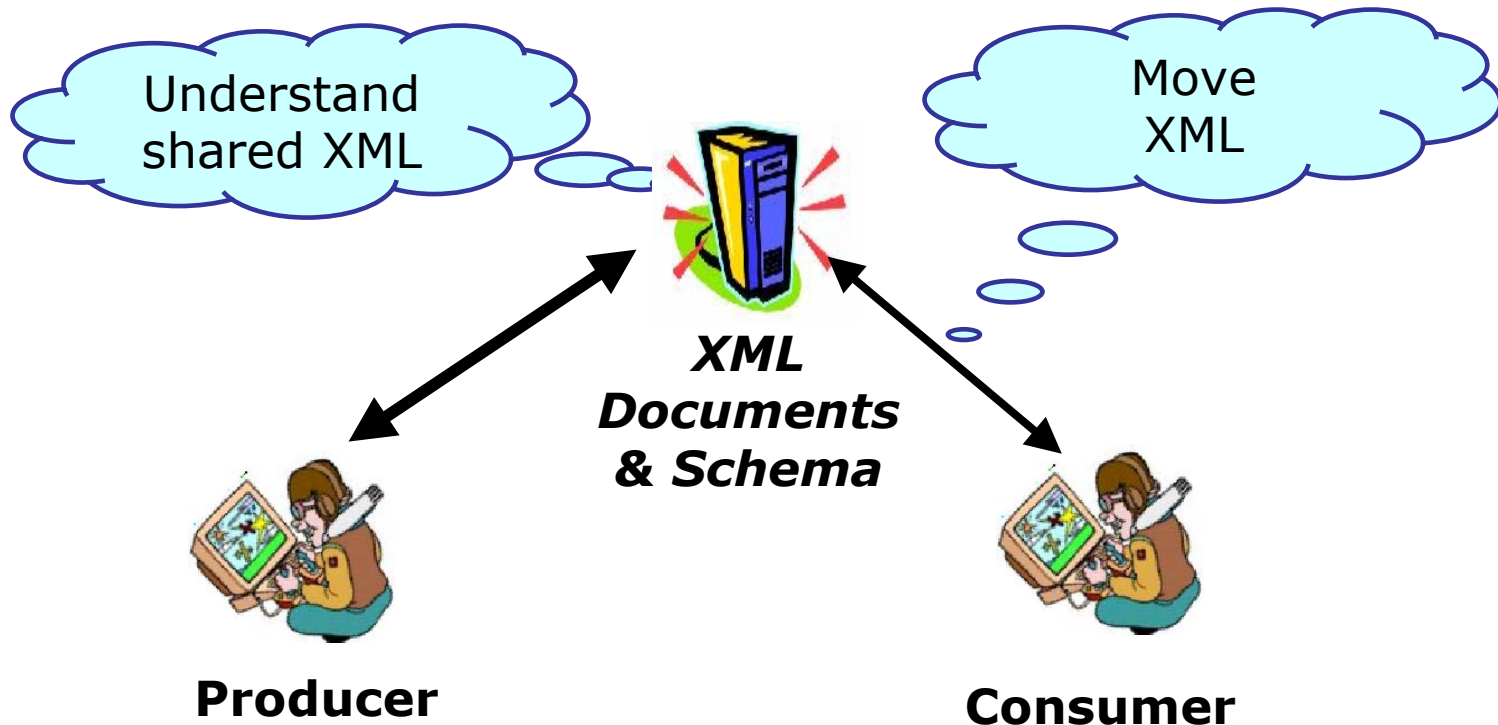
# XML Data-Exchange Illusion



# What we think XML is about

- **Leverage high-performance DBMS for XML**  
Query engine, concurrency control, transaction processing, access control, ...
- **Publishing XML**  
Convert legacy data into XML  
How to specify & evaluate mapping from non-XML data to XML
- **Storing XML**  
Map XML into underlying storage system  
How to “shred” XML data into non-XML data
- **Querying XML**  
Efficient evaluation of queries over virtual XML  
How to translate queries into native query language
- **Hard, but solvable problems**

# XML Data-Exchange Illusion



# What XML is *REALLY* about

## ■ ***Sharing information***

### ***Understanding & processing shared data***

- Tools for data modelling
  - Requires expert understanding of application domain
- Tools for syntactic (representation) & semantic mappings
  - Understanding semantics of schemas
- Old problem with no standard solutions

## ■ ***Moving data***

### ***Efficient communication & processing***

- On-line compression & decompression
- Querying, filtering, routing of (compressed) XML
- Optimize filters based on known schema of XML data