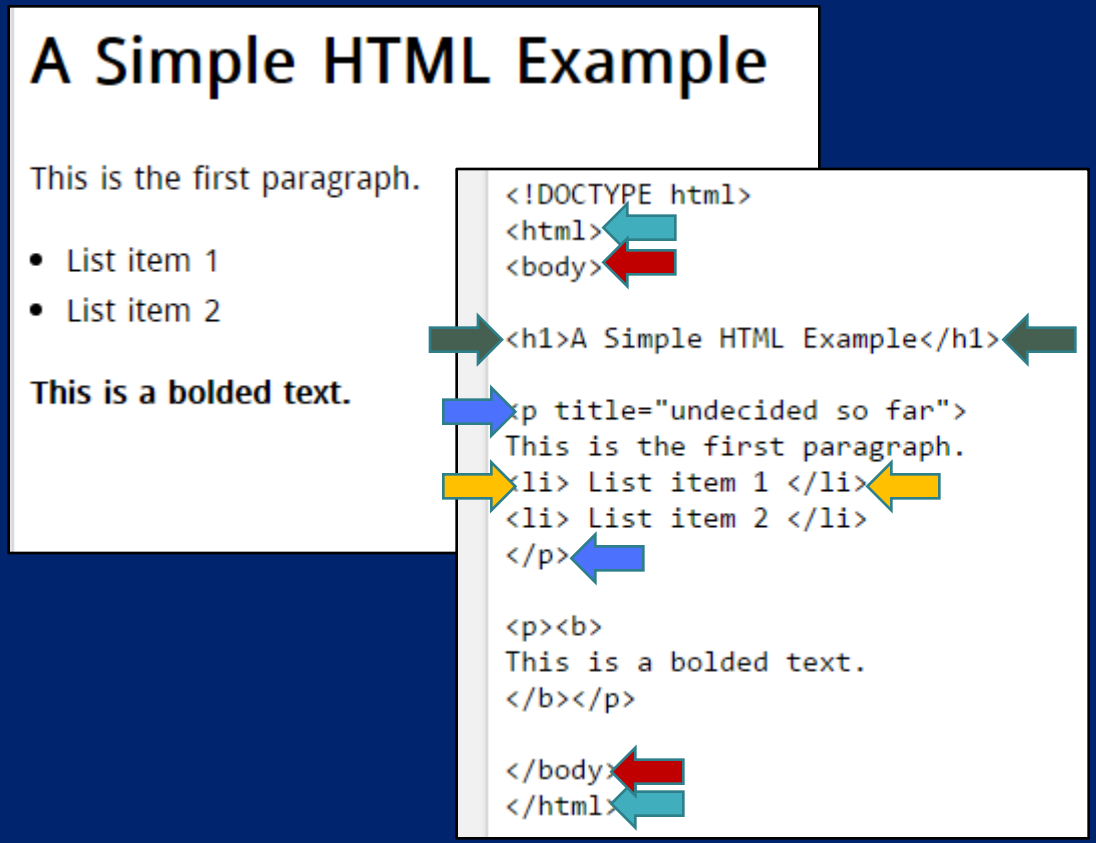
Semi-structured Data Model

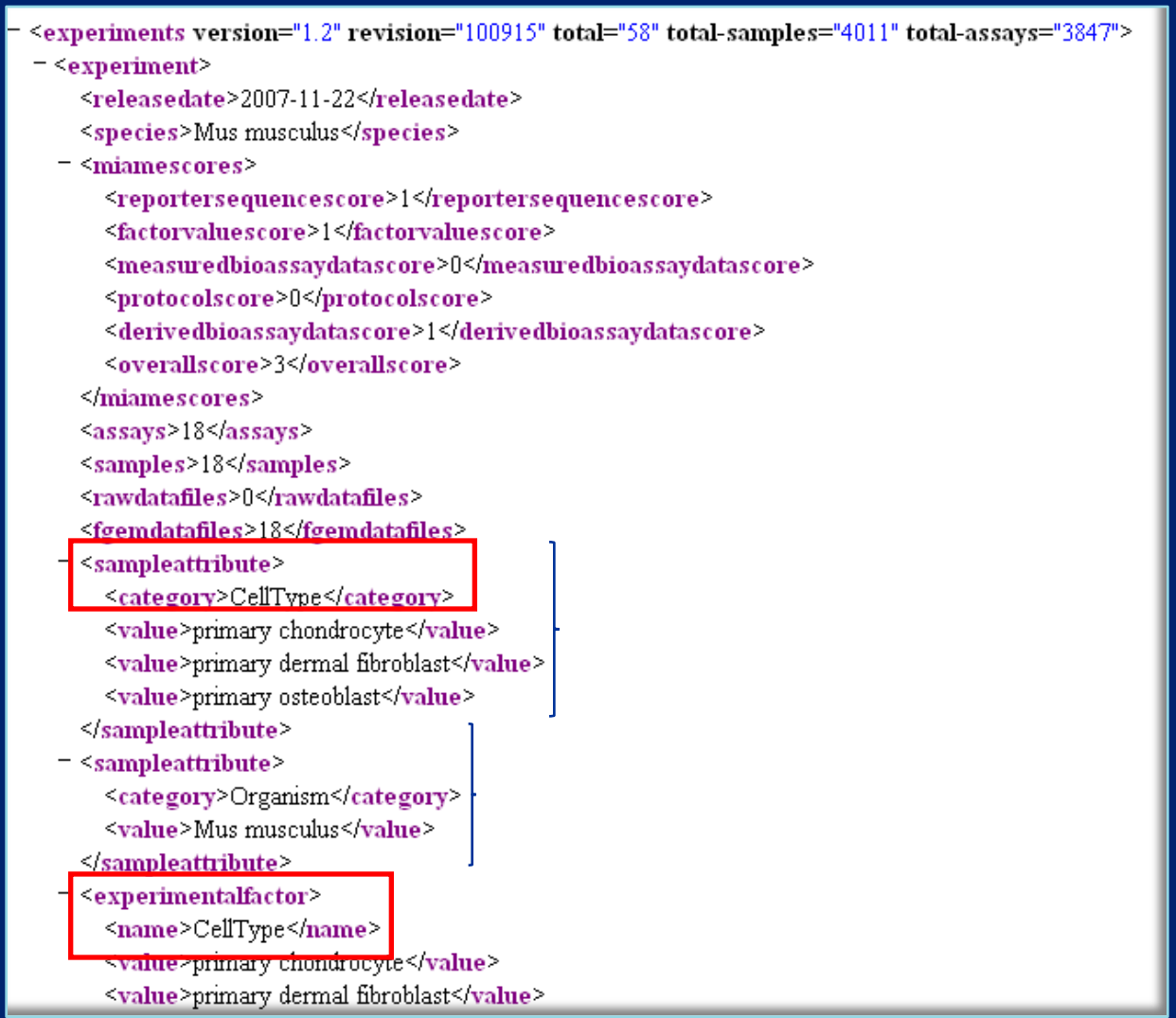
SDSC SAN DIEGO SUPERCOMPUTER CENTER

After this video you will be able to

- Distinguish between structured and "semi-structured model"
- Recognize that most semistructured data are tree-structured
- Explain why tree navigation operations are important for XML, JSON

Let's talk Web Pages

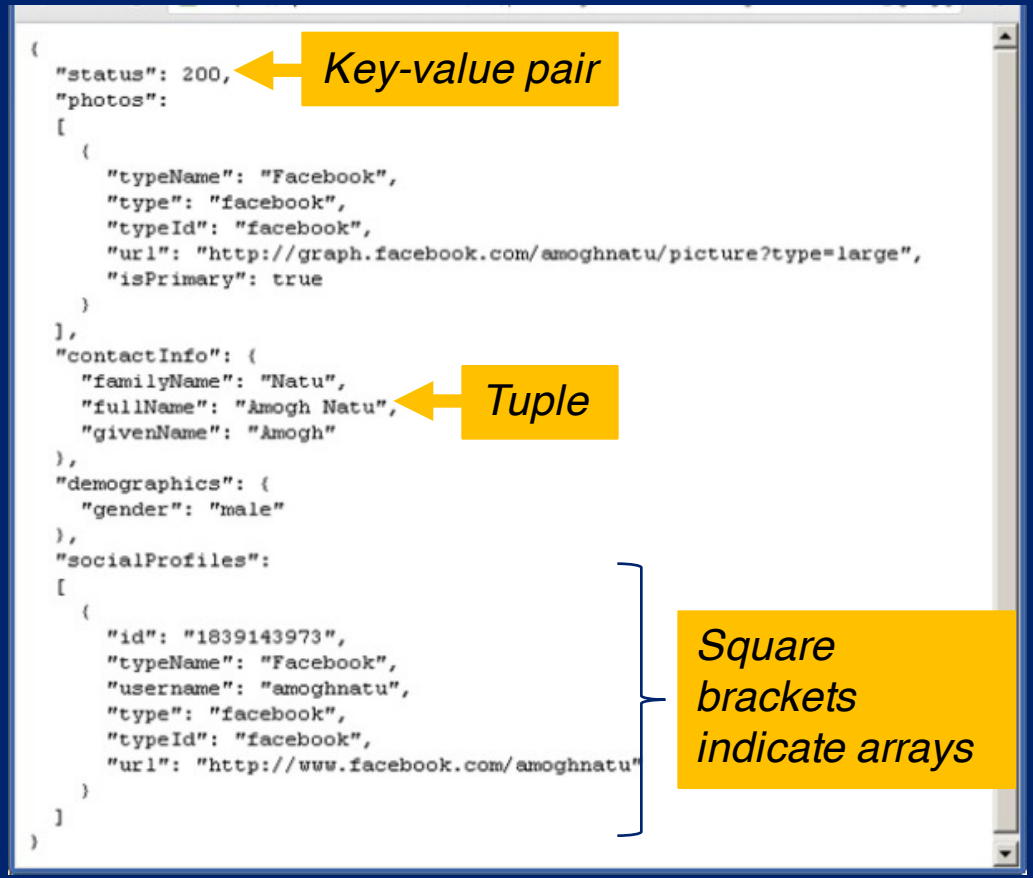




XML

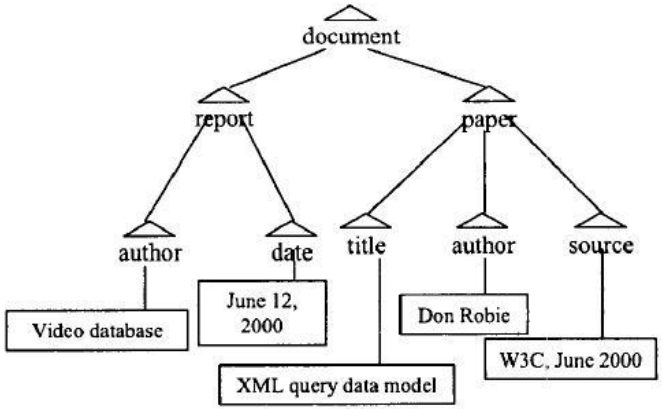
- Allows the querying of both schema and data

JSON


{
"status": 200,
"photos":
[
{
"typeName": "Facebook",
"type": "facebook",
"typeId": "facebook",
"url": "http://graph.facebook.com/amoghnatu/picture?type=large",
"isPrimary": true
}
],
"contactInfo": {
"familyName": "Natu",
"fullName": "Amogh Natu",
"givenName": "Amogh"
},
"demographics": {
"gender": "male"
},
"socialProfiles":
[
{
"id": "1839143973",
"typeName": "Facebook",
"username": "amoghnatu",
"type": "facebook",
"typeId": "facebook",
"url": "http://www.facebook.com/amoghnatu"
}
]
}
Key-value pair
Tuple
Square brackets indicate arrays

Tree Data Structure

```
<document>
  <report>
    <author>Video database</author >
    <date>June 12, 2000</date>
  </report >
  <paper>
    <title>XML query data model</title>
    <author>Don Robie</author>
    <source>W3C, June 2000</source>
  </paper>
</document>
```



Tree Operations

- **Paper**
  - getParent → document
  - getChildren →title, author, source
  - GetSibling → report
- **"Video database"**
  - Root-to-Node path → document/report/author/"video database"
- **Queries need tree navigation**
  - Author of "XML query data model"