putting the pieces together

bringing archival tradition to an item focused web

“Intellectual control means knowing what the pieces of an archival collection are, where they come from, and how they fit together.”

- James M. O’Toole
putting the pieces together
archival tradition?

“The archives cannot be understood without a knowledge of the Administration which produced them.”
- Sir Hilary Jenkinson A Manual of Archival Administration

“Arrangement not only affects the accessibility of archives; it also affects their evidential values, for archives must be arranged in a manner that will protect their integrity as historical evidence.”
- T. R. Schellenberg Modern Archives

“Restoring the coherence that was originally [in a collection] becomes a critical task.”
- James M. O’Toole Understanding Archives and Manuscripts
putting the pieces together
archivists, physical and digital

how do archivists create this coherence? we:

1. collect meaningful materials within a meaningful focus
2. arrange materials according to provenance and original order
3. research collections and author finding aids

but when we digitize on a large scale we:

1. describe at the item-level
2. use systems that provide item-level access only
3. combine all of our digitized material into one system

why? because that’s how the web has been used...
premier collection of w. e. b. du bois material, 170 linear feet

scanning everything, 100,000 individual items

MODS for item-level description

METS for structural metadata

FEDORA repository

Solr for search

Money from the Verizon Foundation and NEH
**goal**: keep the hierarchical relationships from the collection’s arrangement intact.
series view

W.E.B. Du Bois Papers

Series 1. Correspondence, 1877-1963

350 boxes (175 linear ft)
AIS 312

Background
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed aliquam, nibh sed fringilla facilisis, est etiam elementum nunc, sit amet hendrerit magna aenei

Contents
Fusce et sapien tellus, eget convallis luctus. Duis accumsan blandit arcu, vitae convallis orci consetetur sit amet elemetum non, sit amet hendrerit magna aenei

Subseries
Subseries 1a. General correspondence
Subseries 1b. Crisis correspondence

Subjects
African American History
Social movements—United States—20th century

Contributors
Garvey, Marcus
Washington, Booker T.

Genre
Correspondence (letters)
Maps
Letter from Monroe Trotter to W.E.B. Du Bois, 1906 July 15

Informing Du Bois on behalf of the members of the Niagara Movement that perhaps the best approach to confronting Booker T. Washington’s message would be to cause him physical harm. Du Bois notes for response that Trotter is insane.

Place: Boston (Mass.)
Extent: 3p.
Enclosure: Sketch of a clenched fist
Language: English
Subjects: African Americans -- Mental illness -- 20th Century
Freedom Struggle--United States
King, King, 1925-1932
Contributors: Niagara Movement
NAACP
Genre: Correspondence
Location: Special Collections and University Archives, W.E.B. Du Bois
Library, University of Massachusetts Amherst
MS 312
Rights: Fees may apply for reproduction
URL: http://www...etc/number
Identifier: 785699450
putting the pieces together
hierarchy with mods?

*not so easy*...

```xml
<mods:titleInfo>
  <mods:title>A single item</mods:title>
</mods:titleInfo>

<mods:relatedItem type="host">
  <mods:titleInfo>
  </mods:titleInfo>
</mods:relatedItem>

allowable <relatedItem> types:
- preceding
- succeeding
- original
- host
- constituent
```

Special Collections and University Archives, W. E. B. Du Bois Library, University of Massachusetts, Amherst
putting the pieces together
hierarchy with mods?

**even harder...**

```xml
<mods:titleInfo>
  <mods:title>A single item</mods:title>
</mods:titleInfo>

<mods:relatedItem type="host">
  <mods:titleInfo>
    <mods:title>Series 1A Correspondence</mods:title>
  </mods:titleInfo>
</mods:relatedItem>
```
putting the pieces together
hierarchy with mods?

just plain ridiculous...

<mods:titleInfo>
  <mods:title>A single item</mods:title>
</mods:titleInfo>

<mods:relatedItem type="host">
  <mods:titleInfo>
    <mods:title>Series 1A Correspondence</mods:title>
  </mods:titleInfo>
  <mods:relatedItem type="host">
    <mods:titleInfo>
    </mods:titleInfo>
  </mods:relatedItem>
</mods:relatedItem>
putting the pieces together
our solution
putting the pieces together
what is rdf?

rdf defines relationships using triples:

subject → predicate → object

with fedora’s ontology for defining relationships between digital objects:

item99 → isConstituentOf → series1
<rdf:RDF>
  <rdf:Description rdf:about="info:fedora/credo:coll1">
    <fedora:hasConstituent rdf:resource="info:fedora/credo:series1"/>
    <fedora:hasConstituent rdf:resource="info:fedora/credo:series2"/>
  </rdf:Description>

  <rdf:Description rdf:about="info:fedora/credo:series1">
    <fedora:isConstituentOf rdf:resource="info:fedora/credo:coll1"/>
  </rdf:Description>

  <rdf:Description rdf:about="info:fedora/credo:coll1">
    <fedora:isMemberOfCollection rdf:resource="info:fedora/credo:coll1"/>
  </rdf:Description>

  <rdf:Description rdf:about="info:fedora/credo:coll1">
    <fedora:isMemberOfCollection rdf:resource="info:fedora/credo:coll1"/>
  </rdf:Description>
</rdf:RDF>
putting the pieces together
does this approach work?

- is it usable?
- does it usefully leverage archival tradition?
- could we do better?
putting the pieces together
could we do better?

explore the potential of the web
collocate and link data

take advantage of new technologies
rdf, sparql, augmented reality

continue to develop archival standards
eac, ead, rda, ???

incorporate existing outside standards
foaf, relationship, event

1. http://www.w3.org/DesignIssues/LinkedData
2. http://www.w3.org/TR/rdf-primer/
3. http://www.w3.org/TR/rdf-sparql-query/