meta-metadata:
an extensible semantic architecture for multimedia metadata definition, extraction, and presentation

andruid kerne, sashikanth damaraju, bharat kumar, and andrew webb

interface ecology lab

department of computer science & engineering
texas a&m university

information sources

large number of elements

human experience: difficult to explore

metadata semantics

about elements within a collection

used to search, browse, collect and explore

find related information

unified framework

problem: heterogeneous sources

solution: language for authoring
extensible
metadata structure

meta-metadata

language
defines metadata
life cycle
representation
extraction
presentation

```
<meta_metadata name="acm_portal" url_base="http://portal.acm.org/citation.cfm?"
    extends="pdf" package="ecologylab.semantics.library.scholarlyPublication">
    <meta_metadata_field name="title" navigates_to="full_text" scalar_type="String"
        xpath=".//td[@class='medium-text']/strong/child::text()" />
    <meta_metadata_field name="source" comment="Metadata related to where...">
        <meta_metadata_field name="heading" navigates_to="archive" scalar_type="String"
            xpath=".//span[@class='mediumb-text']/child::text()" is_facet="true" />
        <meta_metadata_field name="archive" hide="true" scalar_type="ParsedURL"
            xpath=".//span[@class='mediumb-text']/../a[child::text()='archive']/@href" />
```

authored in XML

programming not required

architecture
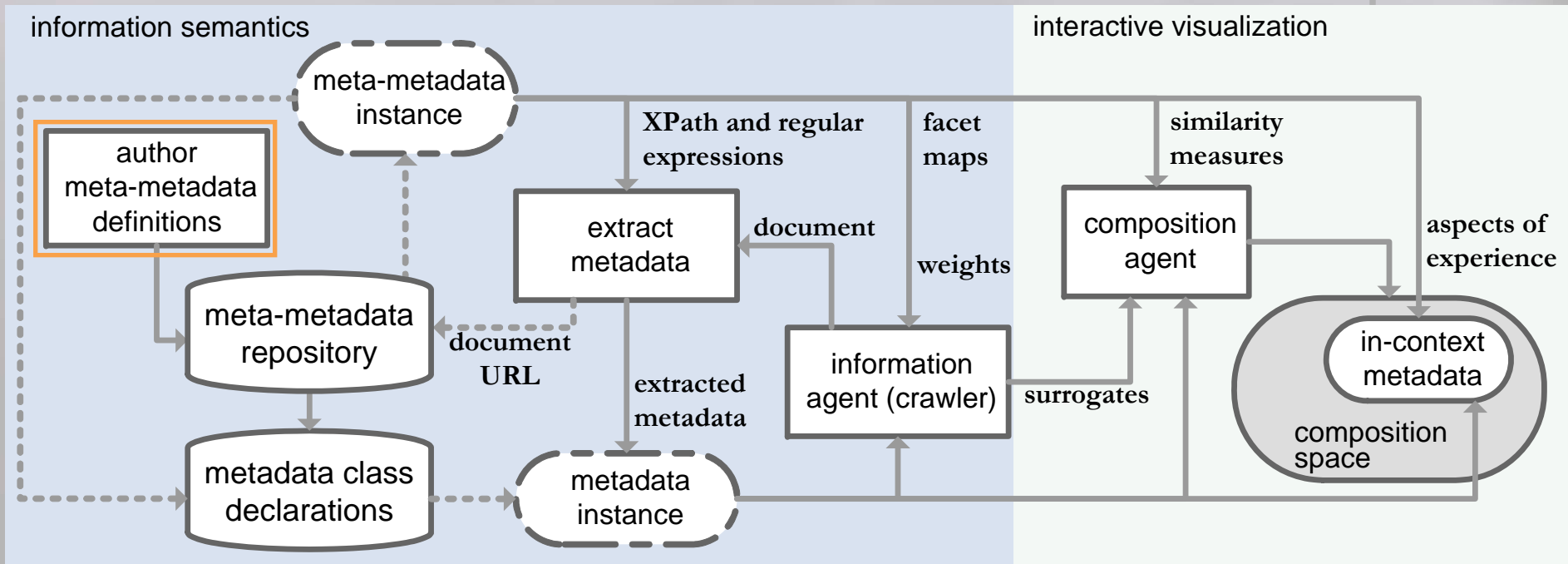
metadata representation
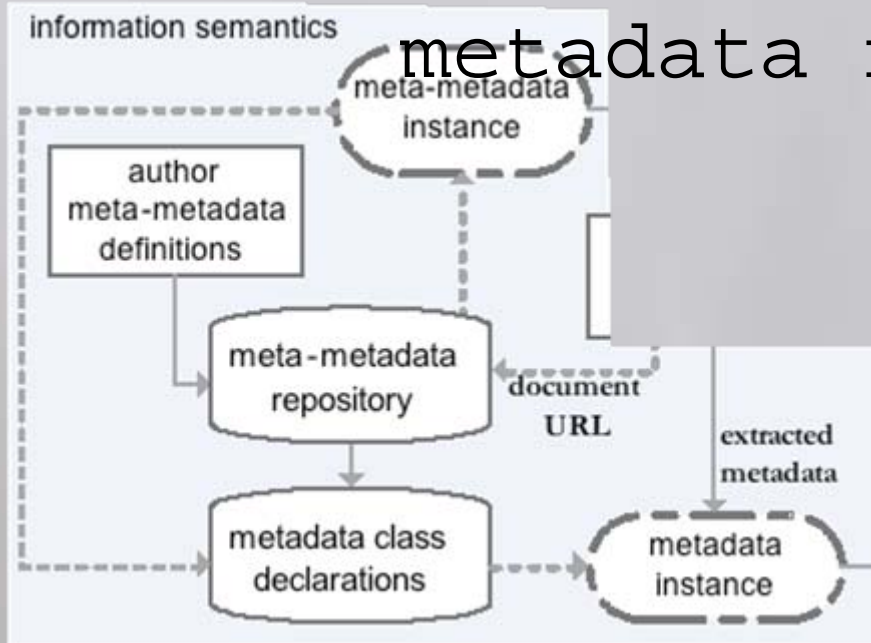
information semantics

meta-metadata instance

author meta-metadata definitions

meta-metadata repository

document URL

metadata class declarations

extracted metadata

metadata instance

```
public class AcmPortal
{
    @xml_nested String title;
    @xml_nested ParsedURL fullText;
    @xml_map HashMapArrayList
        <String, Author> authors;
...
}
```

```
public class Author
{
    @xml_nested String name;
    @xml_nested String affiliation;
    @xml_nested ParsedURL resultsPage;
...
}
```

strongly typed metadata fields

nested structure

meta-metadata > metadata translation

metadata extraction

information semantics

meta-metadata instance

author meta-metadata definitions

XPath and re expressions

extract metadata

doc

meta-metadata repository

document URL

metadata class declarations

extracted metadata

metadata instance
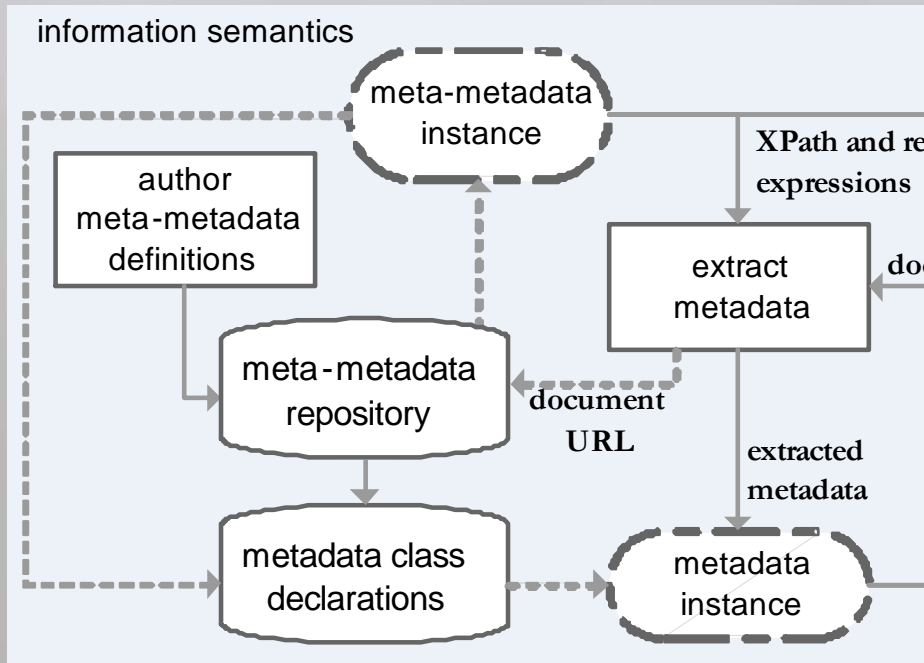
unique by field per a source

XPath expressions

metadata visualization

properties describing
display and interaction
field order
navigatesTo
hide
alwaysShow

combinFormation

mixed-initiative creativity support tool

surrogates & information composition

in-context metadata

location		http://portal.acm.org/ft_gateway.cfm?id=1057063&type=pdf&coll=GUIDE&dl=GUIDE&CFID=10553864&CFTOKE...
title		Breakaway: An Ambient Display Designed to ChangeHuman Behavior

source

heading	Conference on Human Factors in Computing Systems
pages	1945 - 1948
year_of_publication	2005
isbn	:1-59593-002-7

authors (4)

name	Nassim Jafarinaimi
affiliation	Carnegie-Mellon University, Pittsburgh, PA
results_page	http://portal.acm.org/author_page.cfm?id=81100381825&coll=GUIDE&dl=GUIDE&trk=0...

abstract	We present Breakaway, an ambient display that encourages people, whose job requires them to sit for long pe...

references (9)

bib_tex	James Bo Begole , John C. Tang , Rosco Hill, Rhythm modeling, visualizations and ap...

citations (2)

bib_tex	Christoph Obermair , Bernd Ploderer , Wolfgang Reitberger , Manfred Tscheligi, Cues ...

location	

contextual display and editing

support for meta-metadata and nested structure

demo

conclusion

diverse metadata from heterogeneous
sources

meta-metadata defines life cycle of
metadata

author XML, no programming

in-context visualization

developed as reusable components

questions and thoughts


```
<meta_metadata name="acm_portal"
    url_base="http://portal.acm.org/citation.cfm">
    <meta_metadata_field name="title" scalar_type="String"
        xpath=".//td[@class='medium-text']/strong/child::text()"/>
    <meta_metadata_field name="full_text"
        scalar_type="ParsedURL" xpath=".//td[...]"/>
    <meta_metadata_field name="author">
        <meta_metadata_field name="name"
            scalar_type="String" xpath=".//td[...]"/>
        <meta_metadata_field name="affiliation"
            scalar_type="String" xpath=".//td[...]"/>
        <meta_metadata_field name="result_page"
            scalar_type="Pars.." xpath=".//td[...]"/>
    </meta_metadata_field>
    ...
</meta_metadata>
```
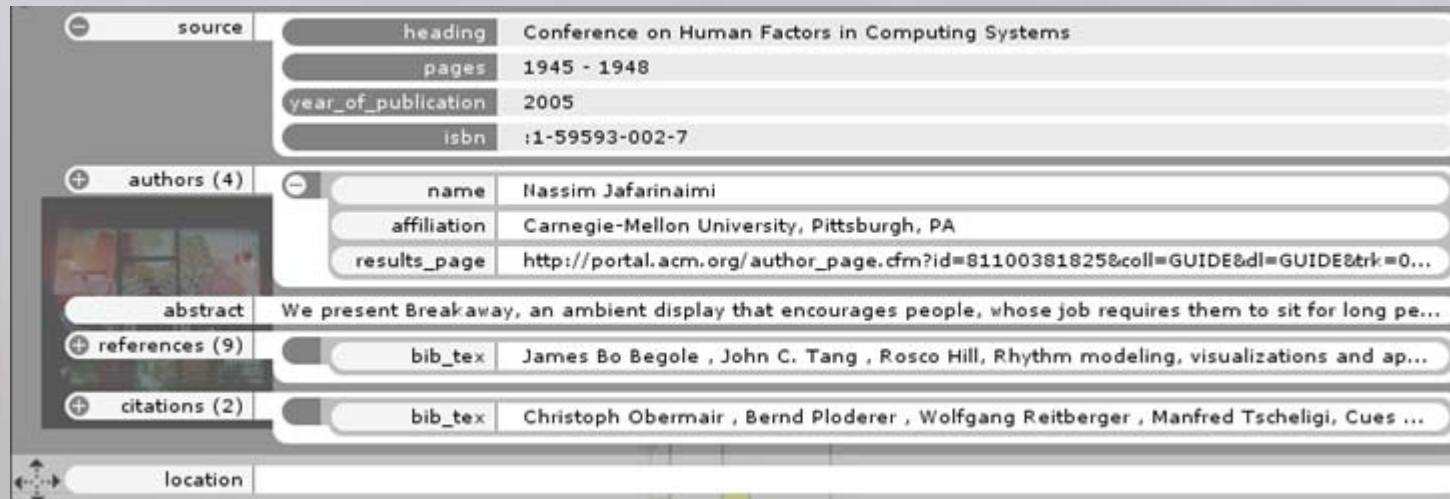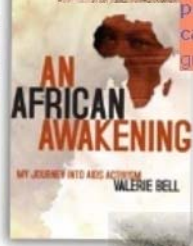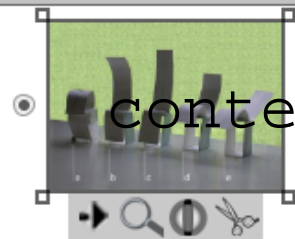
```
public class AcmPortal
{
    @xml_nested String title;
    @xml_nested ParsedURL fullText;
    @xml_map HashMapArrayList
        <String, Author> authors;
    ...
}
```

```
public class Author
{
    @xml_nested String name;
    @xml_nested String affiliation;
    @xml_nested ParsedURL resultsPage;
    ...
}
```

source	
pages	1945 - 1948
year_of_publication	2005
isbn	:1-59593-002-7

authors (4)

name	Nassim Jafarinaimi
affiliation	Carnegie-Mellon University, Pittsburgh, PA
results_page	http://portal.acm.org/author_page.cfm?id=81100381825&coll=GUIDE&dl=GUIDE&trk=0...

abstract — We present Breakaway, an ambient display that encourages people, whose job requires them to sit for long pe...

references (9)

bib_tex — James Bo Begole , John C. Tang , Rosco Hill , Rhythm modeling, visualizations and ap...

citations (2)

bib_tex — Christoph Obermair , Bernd Ploderer , Wolfgang Reitberger , Manfred Tscheligi, Cues ...

location

http://ecologylab.net/beta/cf