Content and Bibliographic Theory

CS 431 Architecture of Web Information Systems

Carl Lagoze
Cornell University – Spring 2005

Acknowledgments

- Herbert Van de Sompel – Los Alamos Research Library
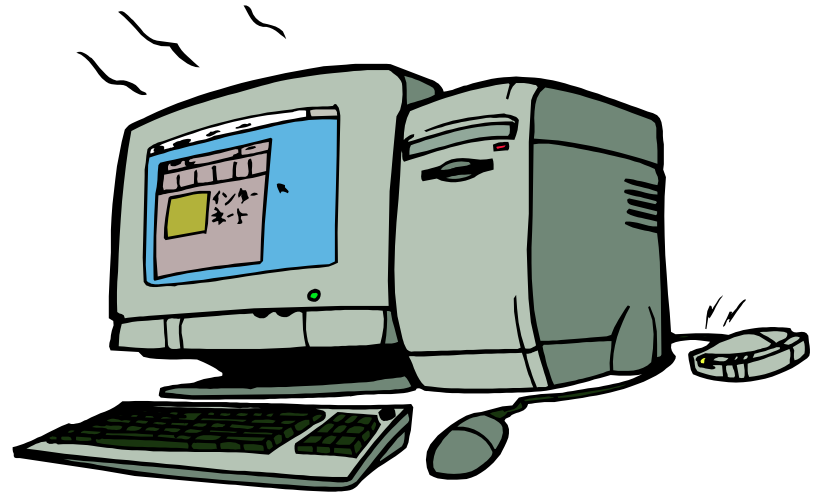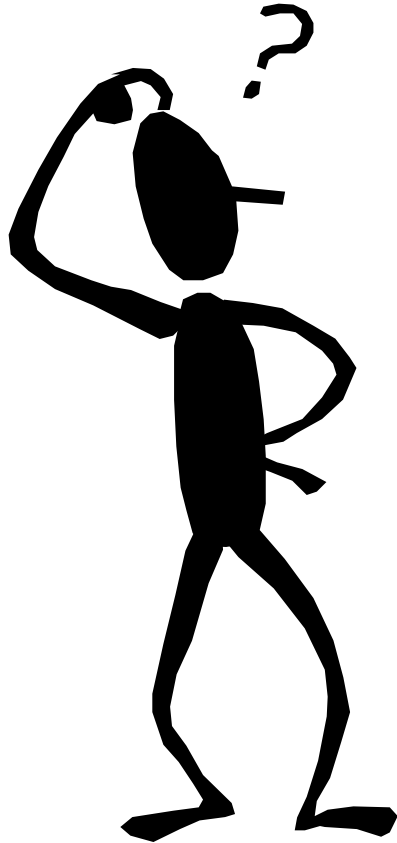- Barbara Tillet - Chief, Cataloging Policy and Support Office Library of Congress
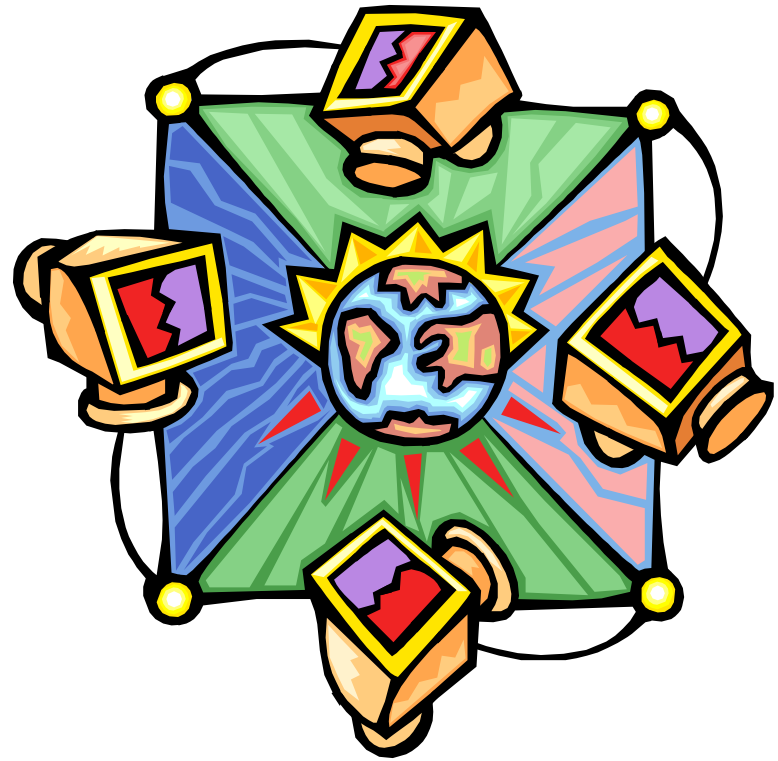
People want stuff.

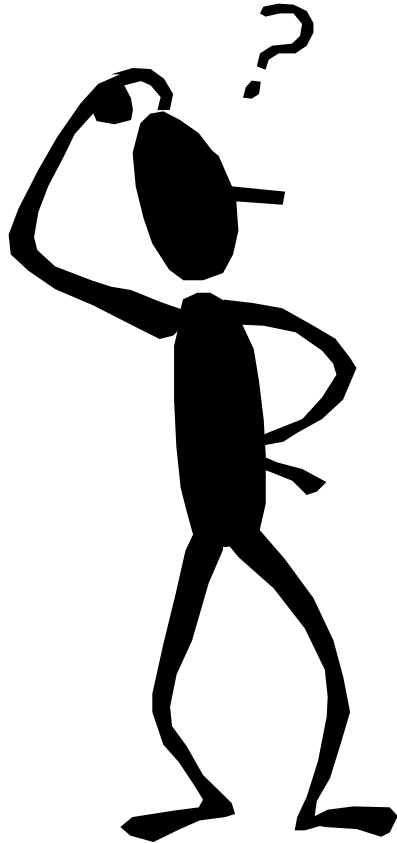Godfrey Rust 1999

Stuff is more complex than you think.

Carl Lagoze 2006

Where did I put that file?

Where is that information?

Am I getting compensated for my talent?

Copies?
Derivations?
Contributions?

Is that available in a way I can use it?

Are there other resources like these?

Information vs. Data – Formal basis

- Claude Shannon – Problem of noisy communication channels
- *Entropy*
  - Informally - measure of the amount of information in a data transmission
  - Amount of disorder in a system.
  - Proportional to the uncertainty of the recipient of a data stream about the content of the message

- Implications
  - Same content can be *encoded* in multiple data streams
  - These different streams mathematically have different information content even though they may be expressions of same information (preservation implications)

Same Information?

24-bit color

8-bit color

Content, Data, Metadata -- **informal definitions**

Content refers to resources as *information* that is of interest to a user. It is the human view of information:

music	*Beethoven's Fifth Symphony*
database	*Genome Database*
literature	*Gone with the wind*
web site	*weather.com*
software	*MS Word*

Content, Data, Metadata -- **informal definitions**

Data emphasizes on the bits and bytes to be processed by a computer. It is the computer representation of information:

        bit and byte layout (e.g., ASCII)
        compression schemes (e.g., MP3)
        image format (GIF, JPEG, PNG)

Content, Data, Metadata -- **informal definitions**

Metadata is data about data/content.

     Descriptive metadata (e.g., catalog records)

     Administrative metadata (e.g. circulation records)

     Structural metadata (e.g. serials record)

   Rights metadata (e.g. shrink wrap license)

Provides a tangible **surrogate** for:

     Abstract information entities

     Available or unavailable physical information entities

Bibliographic model

establishes equivalence classes to organize information objects for human understanding and management

collection

Metadata surrogates:
- Identity
- Description
- Structure

object = piece of content
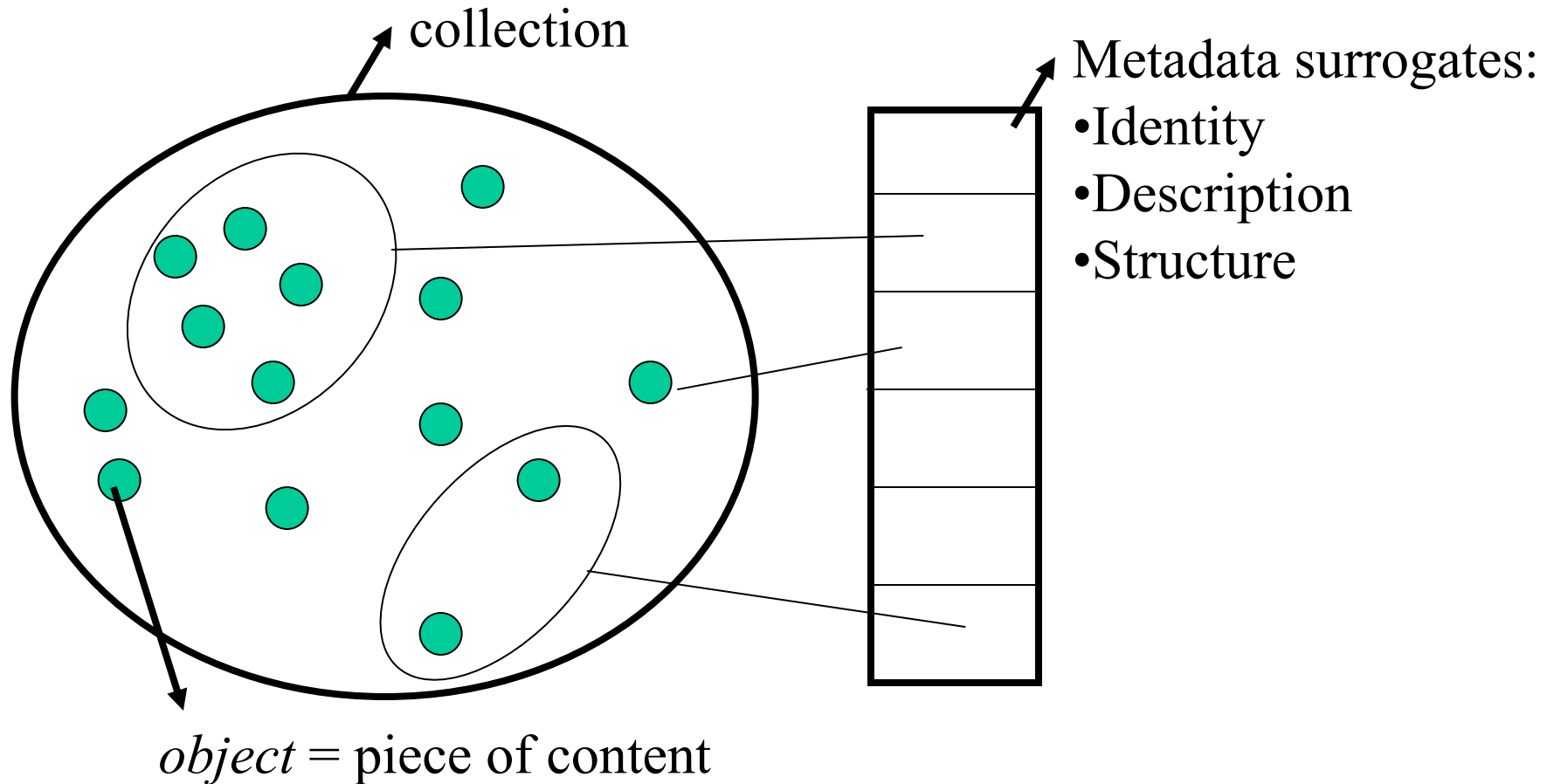
Objectives of a bibliographic system

1. To locate objects in a file or database as the result of a search using attributes or relationships of the objects:
- Known item search -To find a singular object
- Search - To locate sets of objects representing:
  - All objects corresponding to some information unit
  - All objects by a given author
  - All objects about a given author
  - All objects on a given subject
  - All objects published by a given publisher
- These are all access points into the information space

Objectives of a bibliographic system

2. To identify an *object* (i.e. confirm that a described object corresponds to the sought object or distinguish between objects with similar characteristics)

3. To select an *object* that is appropriate to the user's needs/capabilities

4. To obtain access to an *object* (purchase, loan, license, …)

5. To navigate the file or database (browse)
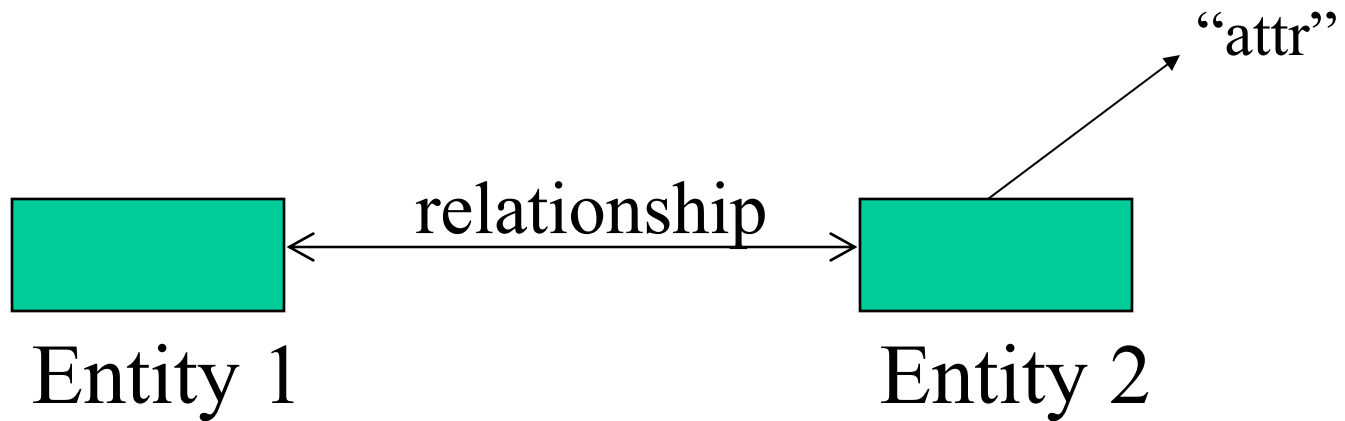
Traditional models challenged by networked digital information

- Scale of corpus or collection
- Variety of content – *Internet Commons*
- Unbinding of information from its carrier
- Mutability/dynamic nature of content
- 'universal context' – all types of people, resources, needs

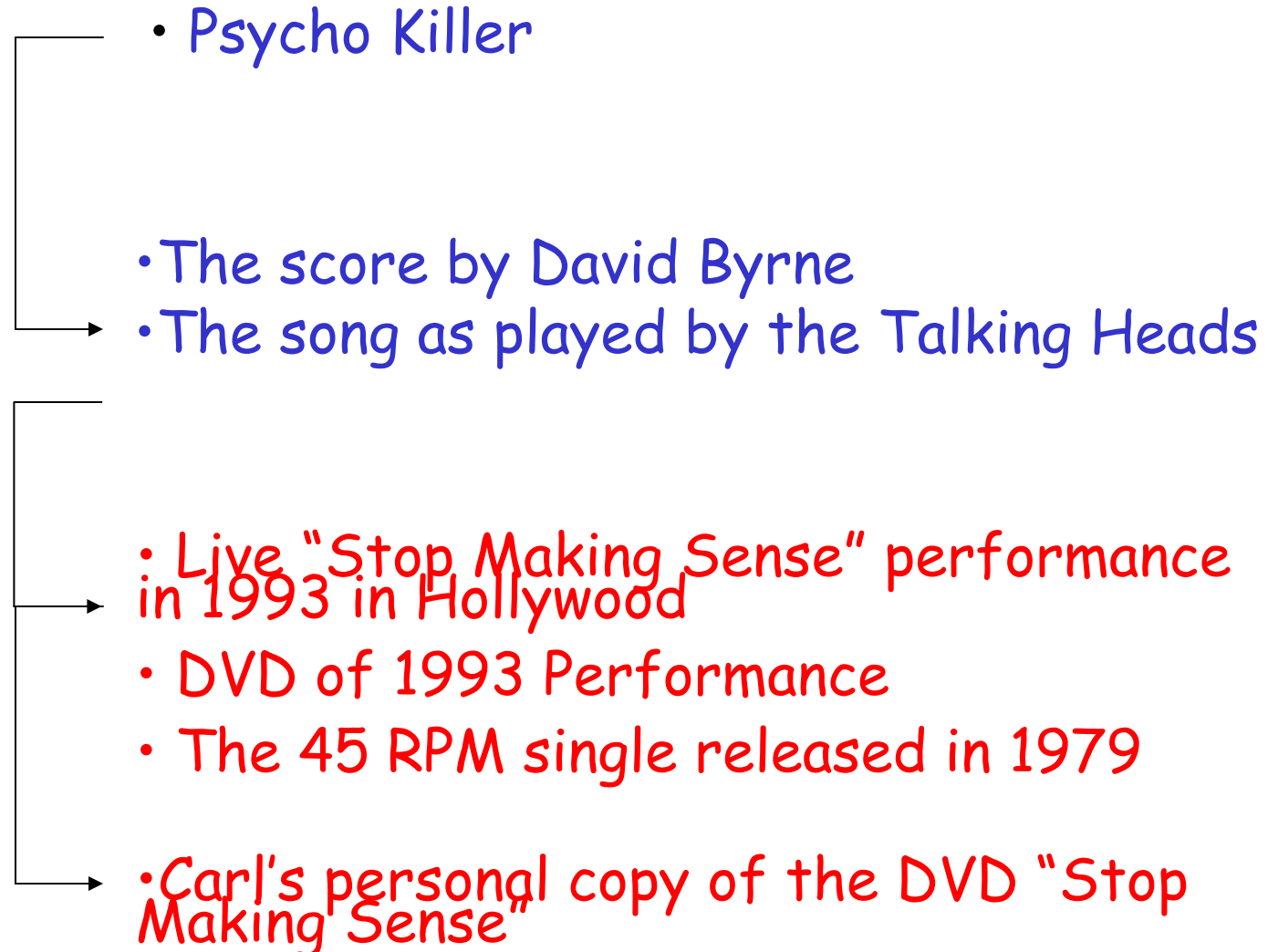Requires more advanced data models to represent:
- Distinct entities
- Their attributes
- Their relationships
- Their evolution over time

Entity-Relationship Model

- Entities
- Relationships
- Attributes

"attr"

Entity 1 ⟷ relationship ⟷ Entity 2

Variants of information entities: Classes with different attributes

- Psycho Killer

- The score by David Byrne
- The song as played by the Talking Heads

- Live "Stop Making Sense" performance in 1993 in Hollywood
- DVD of 1993 Performance
- The 45 RPM single released in 1979

- Carl's personal copy of the DVD "Stop Making Sense"
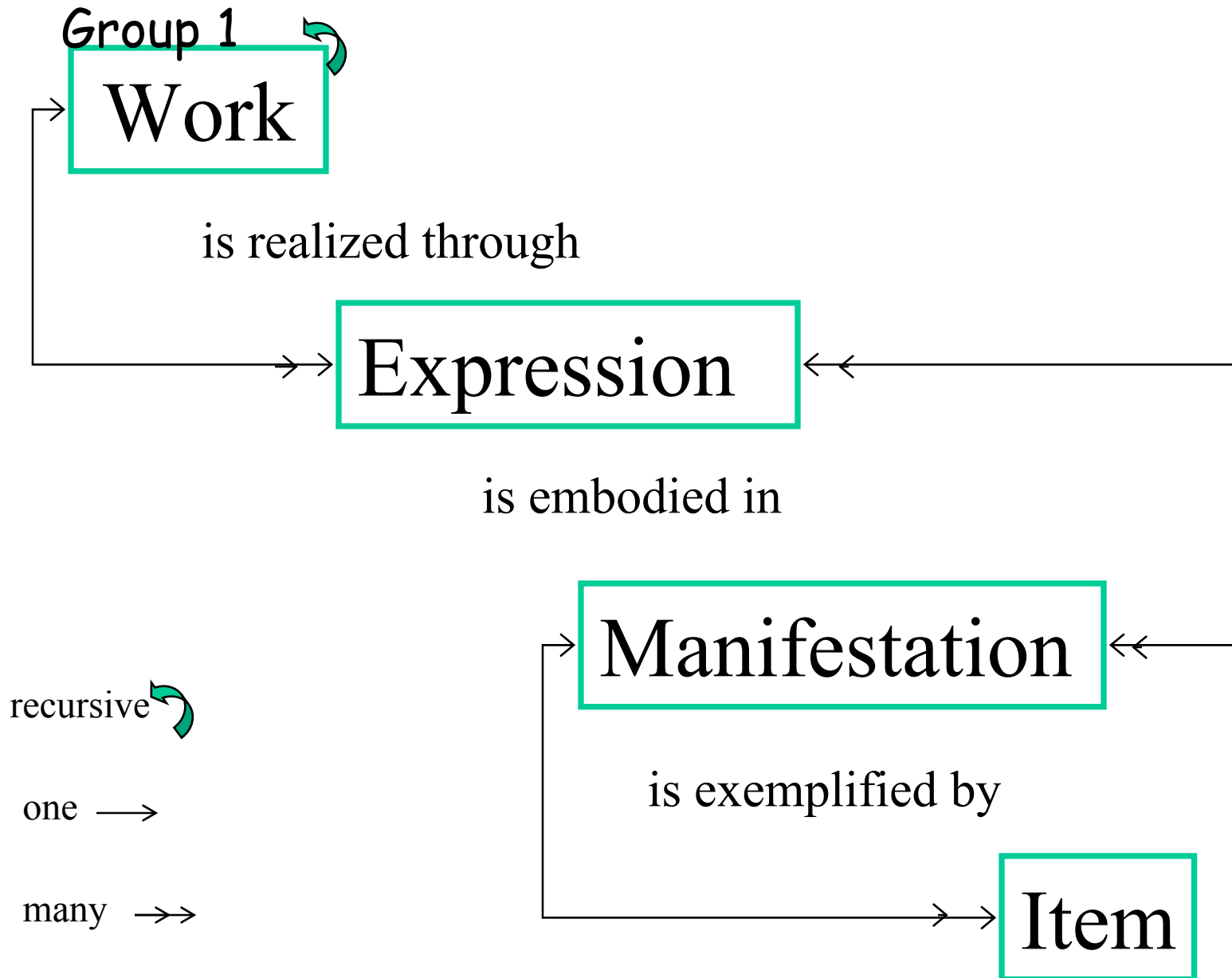
IFLA Model to represent object variants: entities

Entities are the key objects of interest to users of bibliographic data (i.e. of a bibliographic system):

Group 1 - products of intellectual endeavor: work, expression, manifestation, item

Group 2 – the parties responsible for the intellectual content: person, corporate body

Group 3 – the subjects of intellectual endeavor: concept, object, event, place

IFLA model is a conceptual framework.
It does not provide rigorous definitions

Group 1

Work

is realized through

Expression

is embodied in

Manifestation

is exemplified by

recursive

one ⟶

many ⟶⟶

Item

IFLA Model: work, expression, manifestation, item

A work is an **abstract** entity, an idealization e.g.

- The Iliad

- The Weather Channel web site

- Beethoven's Fifth Symphony

- Unix operating system

- The Bible

This is roughly equivalent to the concept of "literary work" used in copyright law.

IFLA Model: work, expression, manifestation, item

An expression is a realization of a work; a representation of the work in a disseminatable form e.g.

- *The Iliad* has oral expressions and written expressions
- A musical work has a score, live performance(s), an original recording, cover(s), ….

Many works have only a single expression, e.g. a web page, or a book which only has a single edition, a painting, a medieval manuscript.

Modes of Expression

- Alpha-numeric
- Music notation
- Other notation (e.g., choreographic)
- Sound
- Still image
- Moving image
- Three-dimensional object
- Combinations of above

IFLA Model: work, expression, manifestation, item

A manifestation is the concrete embodiment of an expression; it reflects <u>physical</u> form e.g.

- The text of *The Iliad* has been manifested in numerous manuscripts and printed books.

- A musical recording can be distributed on CD, cassette, or on a soundtrack of a DVD.

IFLA Model: work, expression, manifestation, item

When many copies are made of a manifestation, each copy is a separate item, e.g.
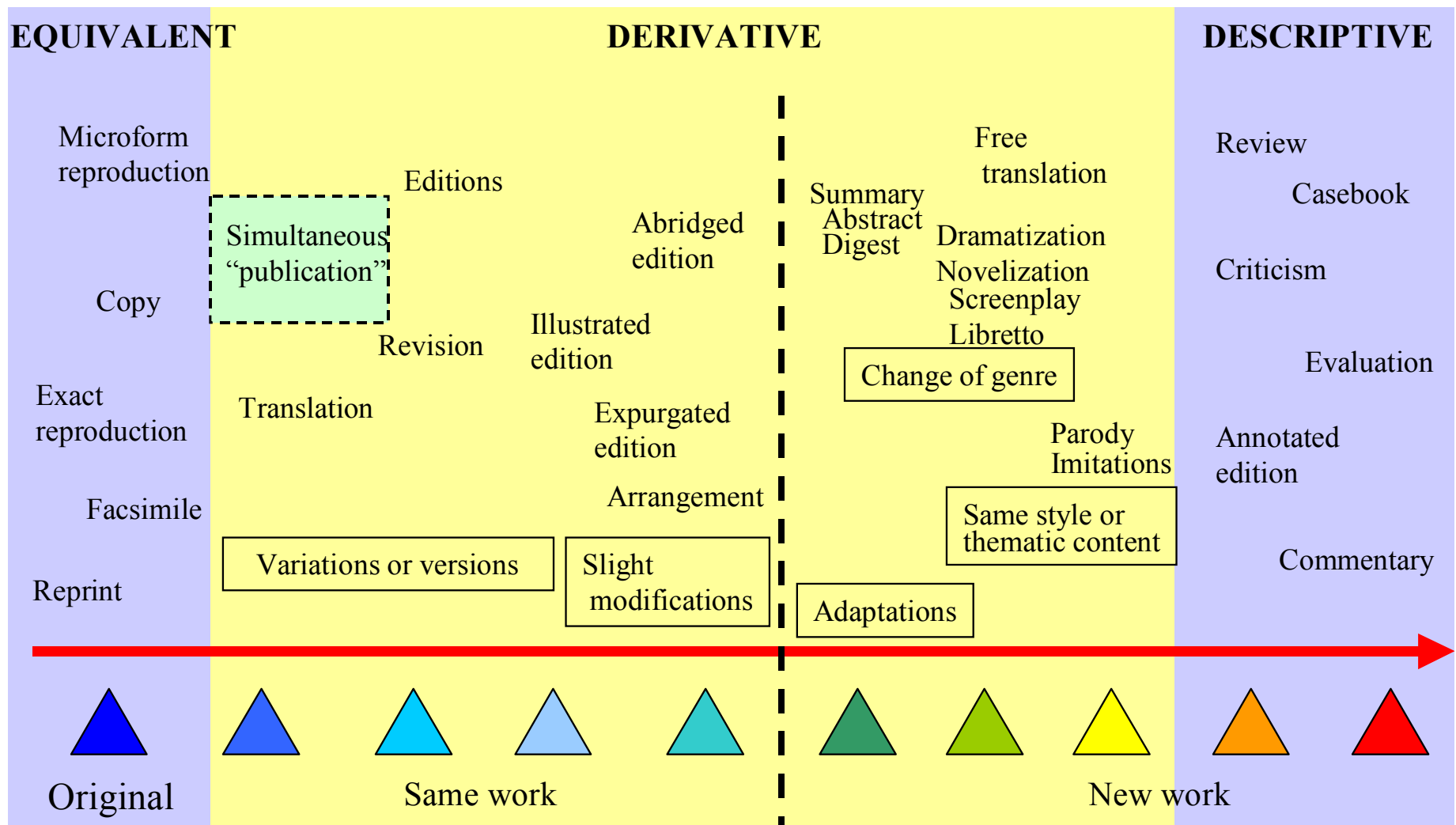
- the Cornell Library's copy of an edition of the Iliad

- your copy of the latest Foo Fighters track

Group 1 Entities' Attributes

FRBR

- **Work**
  - ID
  - Title
  - Date
  - etc.

- **Expression**
  - ID
  - Title
  - Form
  - Date
  - Language
  - etc.

- **Manifestation**
  - ID
  - Title
  - Statement of responsibility
  - Edition
  - Imprint (place, publisher, date)
  - Form/extent of carrier
  - Terms of availability
  - Mode of access
  - etc.

- **Item**
  - ID
  - Provenance
  - Location
  - etc.

Family of Works

EQUIVALENT

Microform reproduction

Copy

Exact reproduction

Facsimile

Reprint

DERIVATIVE

Editions

Simultaneous "publication"

Revision

Translation

Abridged edition

Illustrated edition

Expurgated edition

Arrangement

Variations or versions

Slight modifications

Summary Abstract Digest

Free translation

Dramatization

Novelization Screenplay

Libretto

Change of genre

Parody Imitations

Same style or thematic content

Adaptations

DESCRIPTIVE

Review

Casebook

Criticism

Evaluation

Annotated edition

Commentary

Original

Same work

New work

Cataloging Rules cut-off point

Same Expression

New Expression

New Work

Whole-Part Relationships

- **Components**
- **Aggregates**

Part-to-Part Relationships

- **Component to component**
  - **Sequential Relationship**
  - **Accompanying Relationship (or Companion Relationship)**
    - **Dependent**
    - **Independent**

IFLA Model: work, expression, manifestation, item

CONTENT

work

- Psycho Killer

IFLA Model: work, expression, manifestation, item
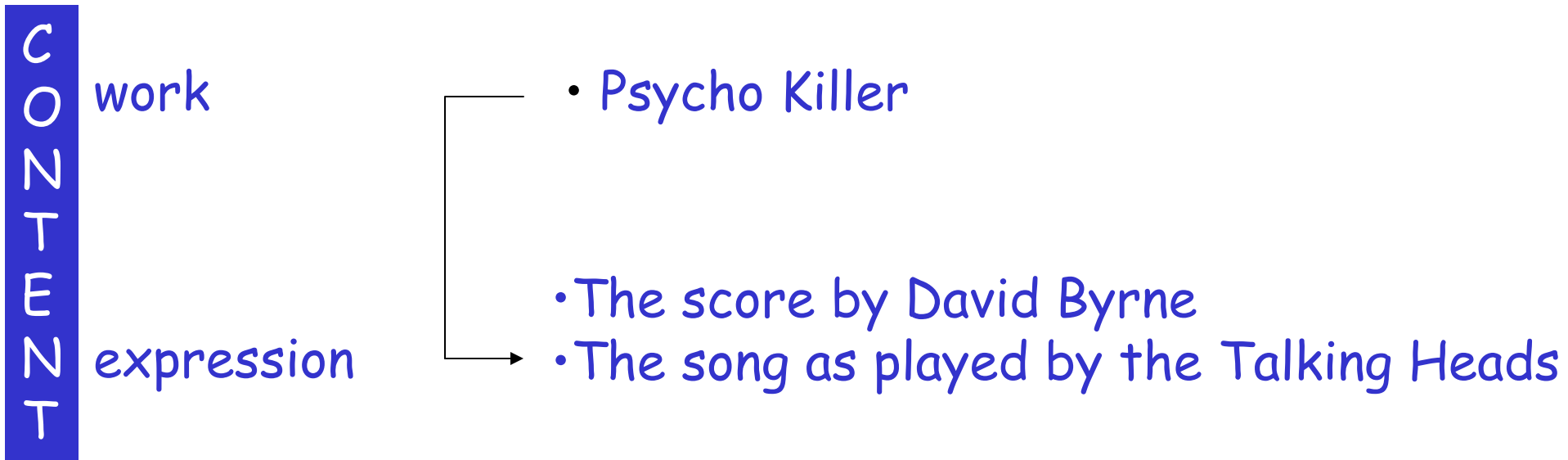
CONTENT

work   • Psycho Killer

expression   • The score by David Byrne
           • The song as played by the Talking Heads

IFLA Model: work, expression, manifestation, item

CONTENT

work
— • Psycho Killer

expression
→ • The score by David Byrne
• The song as played by the Talking Heads

PHYSICAL

manifestation
→ • Live "Stop Making Sense" performance in 1993 in Hollywood
• DVD of 1993 Performance
• The 45 RPM single released in 1979

IFLA Model: work, expression, manifestation, item

C O N T E N T

work — • Psycho Killer

expression →
• The score by David Byrne
• The song as played by the Talking Heads

P H Y S I C A L

manifestation →
• Live "Stop Making Sense" performance in 1993 in Hollywood
• DVD of 1993 Performance
• The 45 RPM single released in 1979

item →
• Carl's personal DVD

IFLA Model: work, expression, manifestation, item

CONTENT

work   • An theory in high energy physics

expression
- A peer-reviewed paper ...
- An oral presentation...
- A preprint ...

PHYSICAL

manifestation
- TeX version posted by the author to arXiv.org
- PDF version created by arXiv.org

item
- The copy of the TeX version on the Italian mirror of arXiv.org
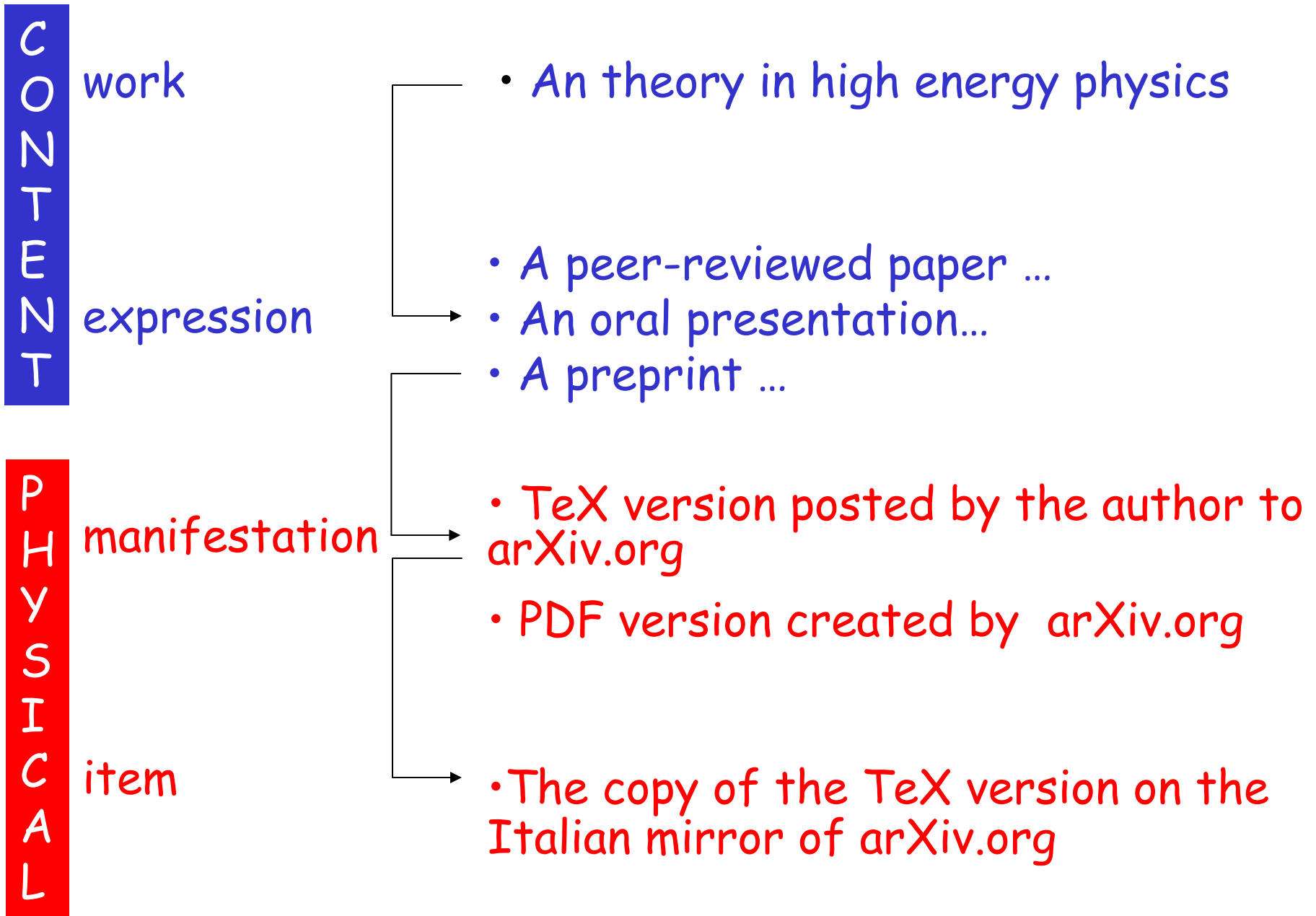
Why should we care about all this?

- Matches a cognitive model of our information seeking and usage behavior
- Impacts intellectual property interests and laws
- Citations and reference linking
- Drives preservation decisions

OCLC WorldCat from Google

Searching for Greece temples

Web   Images   Groups   News   Froogle^{New!}   more »

Google™

greece temples   Search   Advanced Search
Preferences

Web   Results **1 - 10** of about **124,000** for <u>greece</u> <u>temples</u>. (0.10 seconds)

Map 1
darkwing.uoregon.edu/~atlas/europe/static/map08.html - 1k - <u>Cached</u> - <u>Similar pages</u>

Escorted Tours > Escorted Tours: **Greece** > (1) Mainland **Greece** ...
(1) Mainland **Greece**: **Temples** and Mountains. You are equally fascinated by
the beautiful and the historical, by ancient wonder and modern vistas. ...
www.hellenicadventures.com/dynamic/ itinerary.php?itinerary_id=37 - 17k - <u>Cached</u> - <u>Similar pages</u>

Find in a Library: **Greece** : **temples**, tombs, & treasures ...
Greece : **temples**, tombs, & treasures ▪ Publisher: Alexandria, Va. : Time-Life
Books, ©1994. ▪ ISBN: 080949020X 0809490218 ▪ Primary ...
worldcatlibraries.org/ wcpa/ow/6dbc76675b9e71b9a19afeb4da09e526.html - <u>Similar pages</u>

Temples - Great Buildings Online
... **Temple** of Apollo, by Ictinus, at Bassae, **Greece**, -420 to -410. **Temple** ... **Temple**
of Hephaestus, by unknown, at Athens, **Greece**, -449. **Temple** ...
www.greatbuildings.com/types/types/temple.html - 13k - May 10, 2004 - <u>Cached</u> - <u>Similar pages</u>

Maecenas: General Contents
... Go back to top of page. **Greece** and sites outside of **Greece** with Greek architecture. ... Republican
Temples, Santa Maria in Trastevere, Trajan's Market. ...
wings.buffalo.edu/AandL/Maecenas/general_contents.html - 27k - <u>Cached</u> - <u>Similar pages</u>

Ancient Greek Art & Architecture
Greek life was dominated by religion and so it is not surprising that the **temples**
of ancient **Greece** were the biggest and most beautiful.They also had a ...
www.ancient**greece**.com/art/art.htm - 13k - <u>Cached</u> - <u>Similar pages</u>

www.amazon.com/exec/obidos/ASIN/080949020X/
<u>Similar pages</u>

ODYSSEY/**Greece**
Religion was at the center of ancient Greek life, so naturally the **temples** were the
largest and most beautiful buildings. **Temples** also served as civic centers. ...

About WorldCat | Help | For Librarians

WorldCat®
Window to the world's libraries

Greece : temples, tombs, & treasures
- Publisher: Alexandria, Va. : Time-Life Books, ©1994.
- ISBN: 080949020X 0809490218

OCLC LIBR Services:
- OCLC FirstSearch
- Request item
- Connect to the catalog at the OCLC Library & Information Center

Find libraries with the item

Enter a postal code, state, province or country 48910   Go

Find item at ...	City	State/Province Country	Map	Library Info	Library Type
Ann Arbor Public Library	Ann Arbor	Michigan		ⓘ	Public
Baker College Library	Flint	Michigan	🗺		Academic
Genesee District Library	Flint	Michigan	🗺	ⓘ	Public
Lansing Community College	Lansing	Michigan	🗺	ⓘ	Academic
Public Library of Saginaw	Saginaw	Michigan			Public
Other regional libraries ...	**City**	**State/Province Country**	**Map**	**Library Info**	**Library Type**
Baldwin Public Library	Birmingham	Michigan	🗺	ⓘ	Public
Bay County Library Syst	Bay City	Michigan			Public
Bay de Noc Community College	Escanaba	Michigan	🗺	ⓘ	Academic
Cranbrook School	Bloomfield Hills	Michigan	🗺	ⓘ	School
Detroit Public Library	Detroit	Michigan		ⓘ	Public
Farmington Community Library	Farmington Hills	Michigan	🗺	ⓘ	Public

WorldCat
Window to the world's libraries

Greece : temples, tombs, & treasures
• ISBN: 080949020X 0809490218

↵ Return

Lansing Community College Library Catalog

| Previous Record | Next Record | Start Over | MARC display | Export | Request This Title | AboutThisItem |

[ISN ▼] [080949020X] [Entire Catalog ▼] [System Sorted ▼] [Search]

Title	**Greece : temples, tombs, and treasures / by the editors of Time-Life Books.**
Pub info	Alexandria, Va. : Time-Life Books, 1994.

LOCATION	CALL NUMBER	STATUS
TLC 3rd Floor	DF 221.2 .G75 1994	AVAILABLE

Phys descr	168 p. : ill. ; 29 cm.
Series	Lost civilizations
Subjects	Excavations (Archaeology) -- Greece.
	Show similar items
	Greece -- History -- To 146 B.C.
	Show similar items
	Greece -- Antiquities.
	Show similar items
Add author	Time-Life Books.
Notes	Includes index.

| Previous Record | Next Record | Start Over | MARC display | Export | Request This Title | AboutThisItem |

LCC Library Home

A Project of OCLC Research

OCLC Online Computer Library Center   Exit

FictionFinder

Search   Browse   Categories

Search

Search   Clear

Search for: melbourne

in: Settings

AND   mystery

in: Fiction Categories

AND

in: Fiction Categories

AND

in: Subjects

Search   Clear

1   5 Works were found.

1.	Murder On the Ballarat Train	Vers:3	★★★	Greenwood, Kerry
2.	Murder On the Ballarat Train: A Phryne Fisher Mystery	Vers:1	★	Greenwood, Kerry
3.	The Brush-off	Vers:4	★★★★★	Maloney, Shane,
4.	Cat Catcher	Vers:3	★★★	Shaw, Caroline
5.	The Great Melbourne Cup Mystery	Vers:3	★★★	Upfield, Arthur William,

1   5 Works were found.

2.6+ million fiction records from Worldcat, clustered by OCLC's FRBR algorithm

Make greater use of data (genres, settings, imaginary characters, etc)

Work display

http://fictionfinder.oclc.org/WebZ/Authorize?sessionid=0   content directions

| Search | Browse | Categories |

Search

| Search | Clear |

Search for: anna karenina

in: Exact Title

AND

in: Author

AND

in: Fiction Categories

AND

in: Subjects

| Search | Clear |

Title: **"anna karenina"**

< 1 >   4 Works were found.

1.	Anna Karenina		Vers:1	★ ★ ★	Collins, Anne
2.	Anna Karenina: Roman		Vers:1	★ ★	Nikolaev, Lev,
3.	Anna Karenina		Vers:391	★ ★ ★ ★ ★	Tolstoy, Leo,
4.	Anna Karenina: En Precis		Vers:1	★	Wren, A. R.

< 1 >   4 Works were found.

Work/expression display

Work/expression/manifestation

A Project of OCLC Research
OCLC Online Computer Library Center   Exit

FictionFinder

Search   Browse   Categories

Search

Search   Clear

Search for: anna karenina
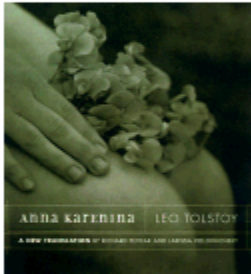
in: Exact Title

AND

in: Author

AND

AND

Russia — Social life and customs — 1533-1917

Manifestation Record Display - Mozilla Firebird

Anna Karenina : A Novel In Eight Parts / Leo Tolstoy ; translated by Richard Pevear and Larissa Volokhonsky.

Edition	1st American ed.
Date	2001.
Publisher	New York : Viking,
Description	xxi, 837 p. ; 24 cm.
ISBN	0670894788
OCLC #	44405534

Literary Fe
Special Fe

Anna Karenina. 198

Manifestation Record Display - Mozilla Firebird

Anna Karenina / Leo Tolstoy ; translated by Louise and Aylmer Maude, with an introduction and notes by W. Gareth Jones.

Notes	Oxford world's classics
Date	1998, 1995.
Publisher	Oxford : Oxford University Press,
Description	xxxiv, 831 p. ; 23 cm.
ISBN	0585387605 (electronic bk.)
URL	http://www.netLibrary.com/urlapi.asp?action=summary&v=1&bookid=56534
OCLC #	49569249

★★★★ (149)
★★★★ (135)   eBooks
★★★★ (131)
★★★★ (123)
★★★★ (115)
★★★★ (113)
★★★★ (112)