College of Information Studies
University of Maryland Hornbake Library Building College Park, MD 20742-4345

Information Infrastructures

Week 1

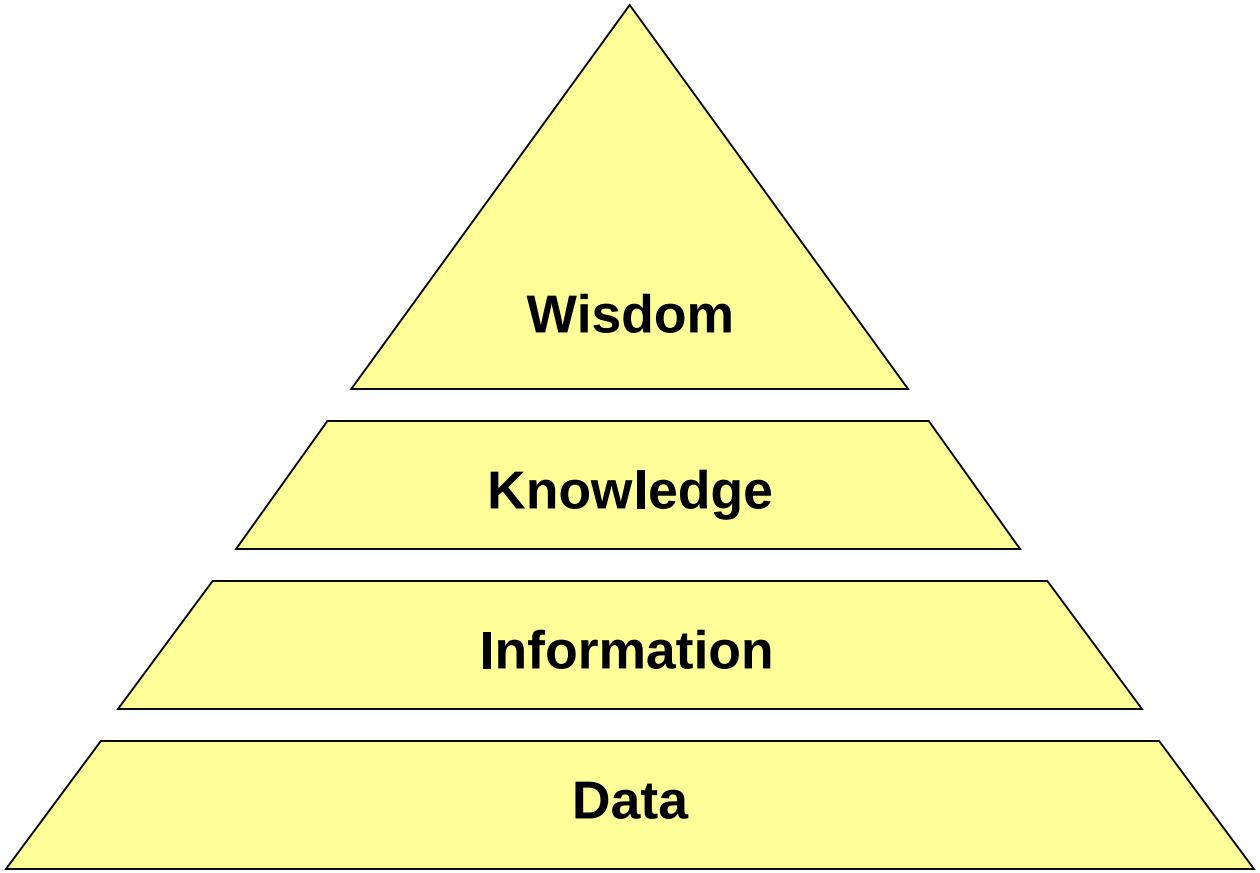
LBSC 671

Creating Information Infrastructures

Tonight

- What's this class about?
- Pieces of the puzzle
- All the usual stuff (syllabus, grading, …)

“Information”



Infrastructures

- Setting
  - Embedded
  - Pervasive
- Design
  - Reflects practice
  - Reflects standards
  - Path dependent
- Learned
- Transparent when it works
  - But visible when it fails!

Star and Ruhleder (1996)

“Information Institutions”

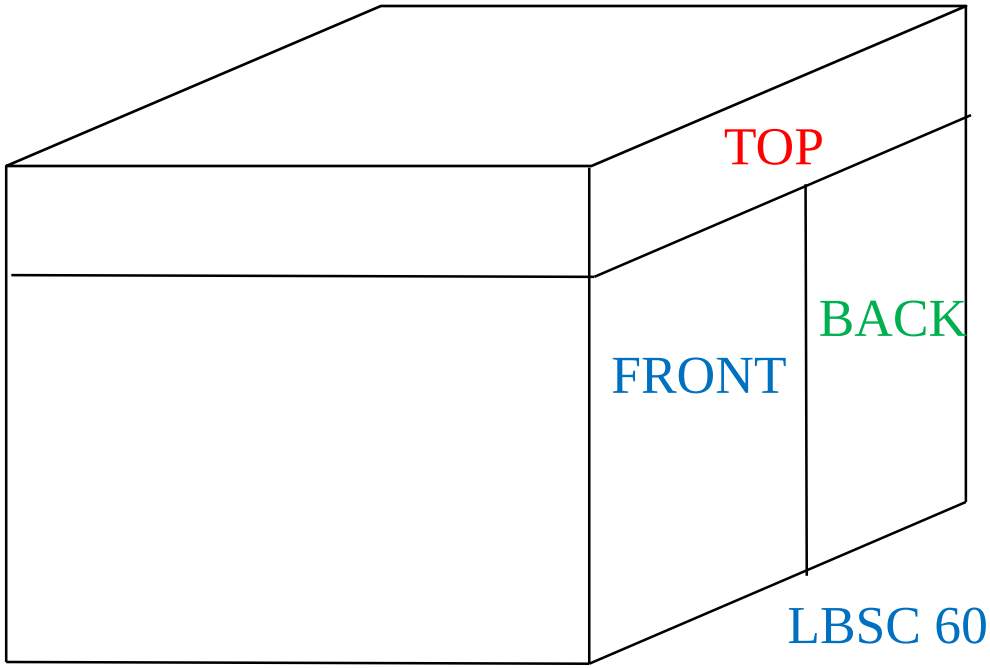
- Schools
  - School libraries, Academic libraries

“LAM”

- Libraries
  - Public, Special, Subscription, (bookstores?)
- Archives
  - Government, Corporate, …
  - Special collections, Historical societies
- Museums
  - Art, Material culture, Natural history

- Search engines
  - Google, Bing, Baidu, Yandex

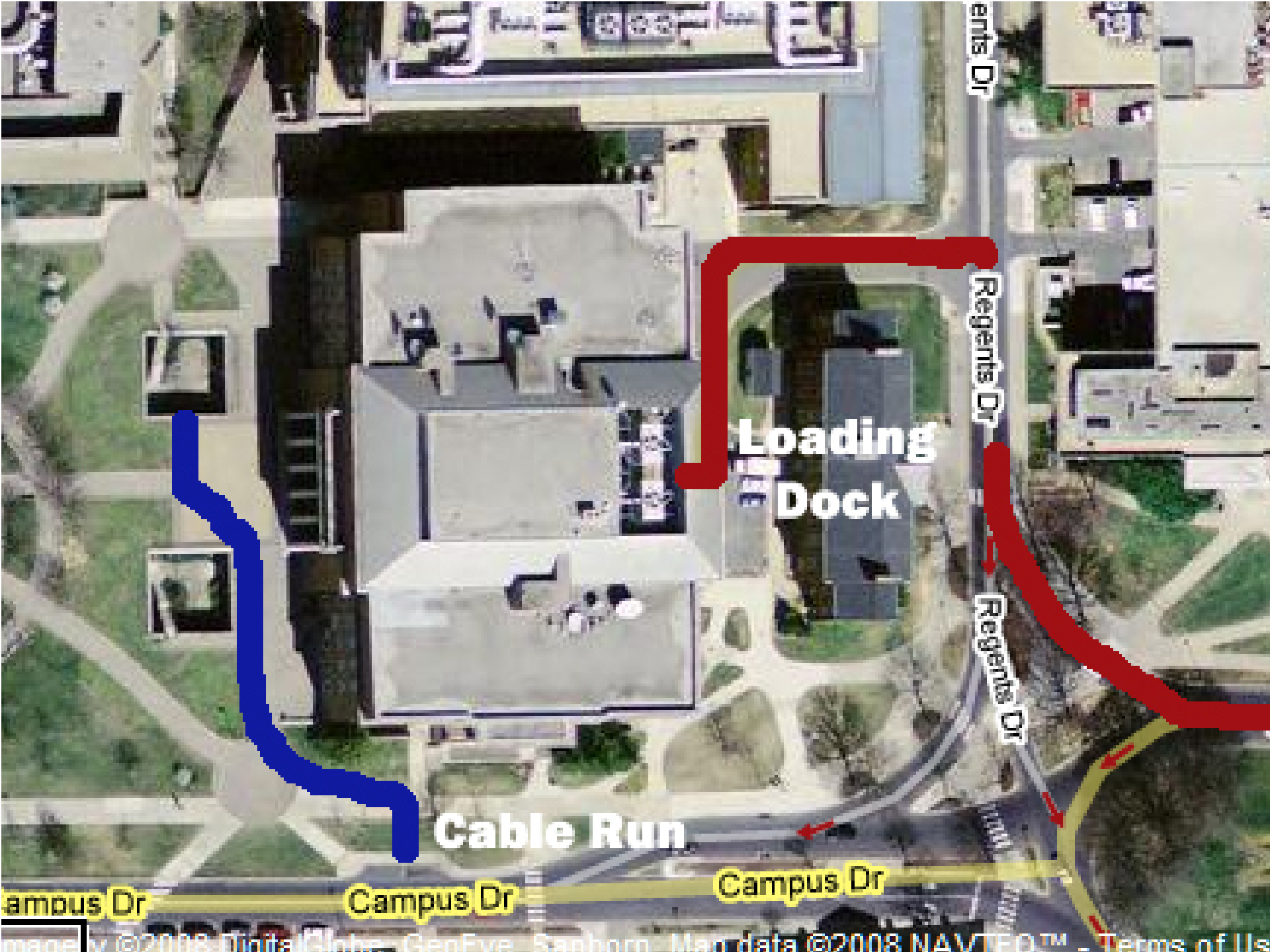
LBSC 791
Designing Principled Inquiry



LBSC 631
Achieving Organizational Excellence

LBSC 671
Creating Information Infrastructures

LBSC 602
Serving Information Needs


Loading Dock
Cable Run
Regents Dr
Regents Dr
Campus Dr
Campus Dr
Campus Dr

The MLS Program

- Specializations
  - Information and Diverse Populations
  - Government Information Management & Services
  - School Library
  - Archives, Records, and Information Management
    - History and Library Science dual degree program
  - Curation and Management of Digital Assets
- Unspecializations
  - Research (Masters thesis option)
  - Individualized Program Plan

Some Related iSchool Courses

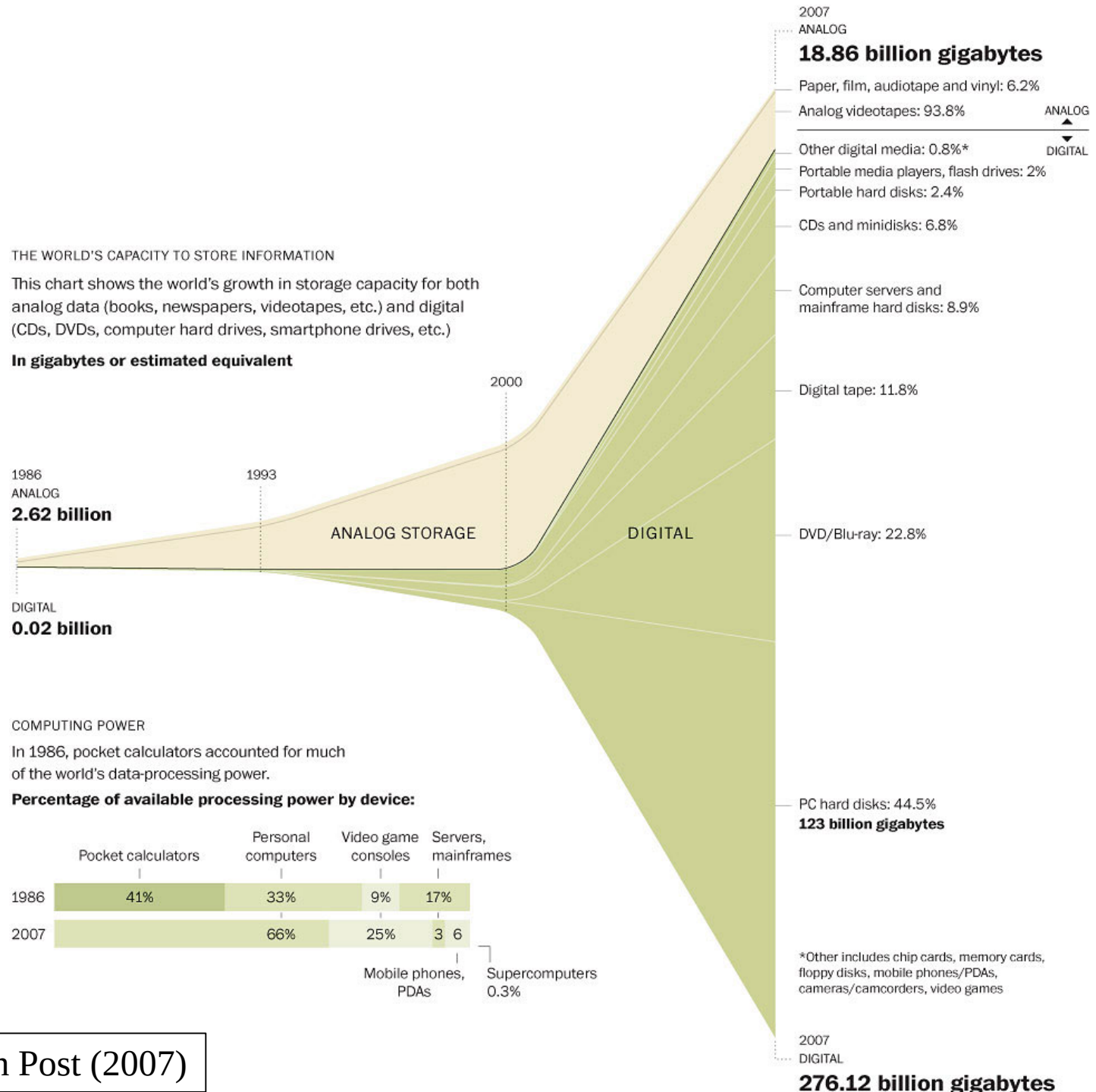
- LBSC
  - 683: Electronic Records
  - 684: Arrangement & Description
  - 708X: E-Discovery
  - 770: Bibliographic Control
  - 773: Classification Theory
  - 783: Technical Services
  - 784: Digital Preservation
  - 785: Appraisal
  - 786: Preservation
- INST
  - 630: Programming
  - 631: HCI Fundamentals
  - 640: Digital Curation Principles
  - 715: Knowledge Management
  - 733: Database Design
  - 734: Information Retrieval
  - 735: Computational Linguistics
  - 737: Digging into Data
- INFM
  - 741: Social Computing
  - 743: Internet Applications

Tonight

- What's this class about?

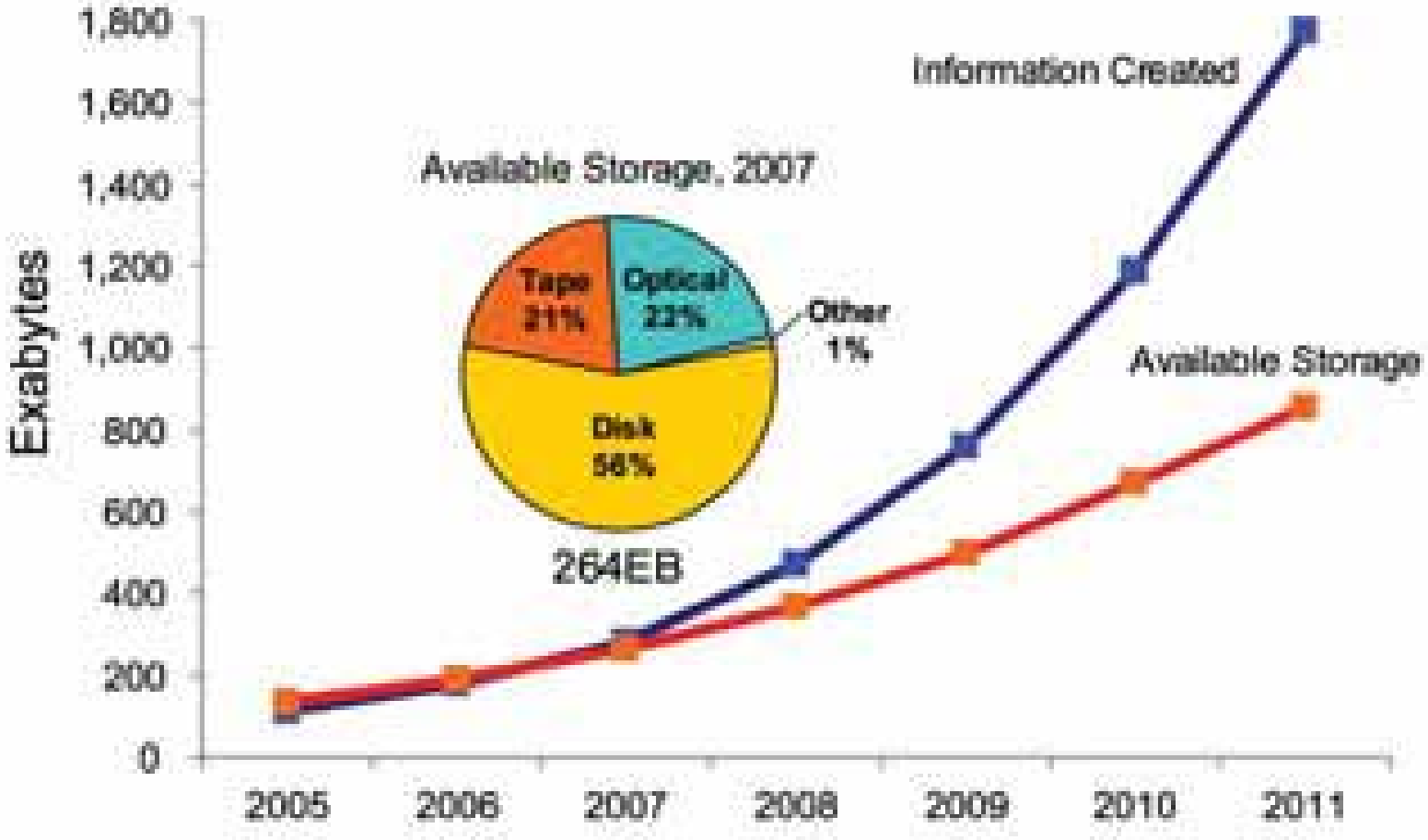
➢ Pieces of the puzzle

- All the usual stuff (syllabus, grading, …)

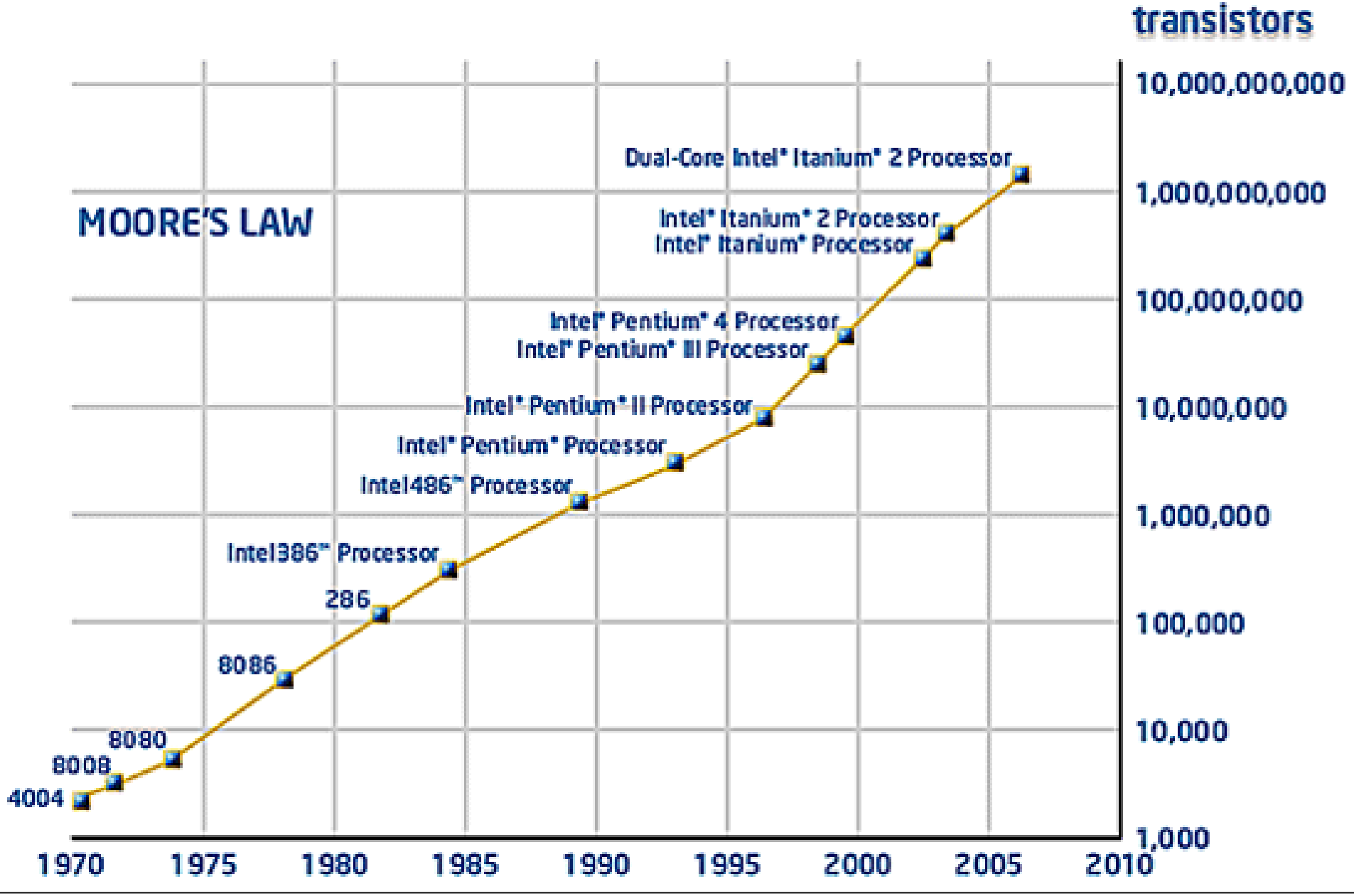


Washington Post (2007)

“Data In Motion” vs. “Data At Rest”



IDC Digital Universe white paper (2008)

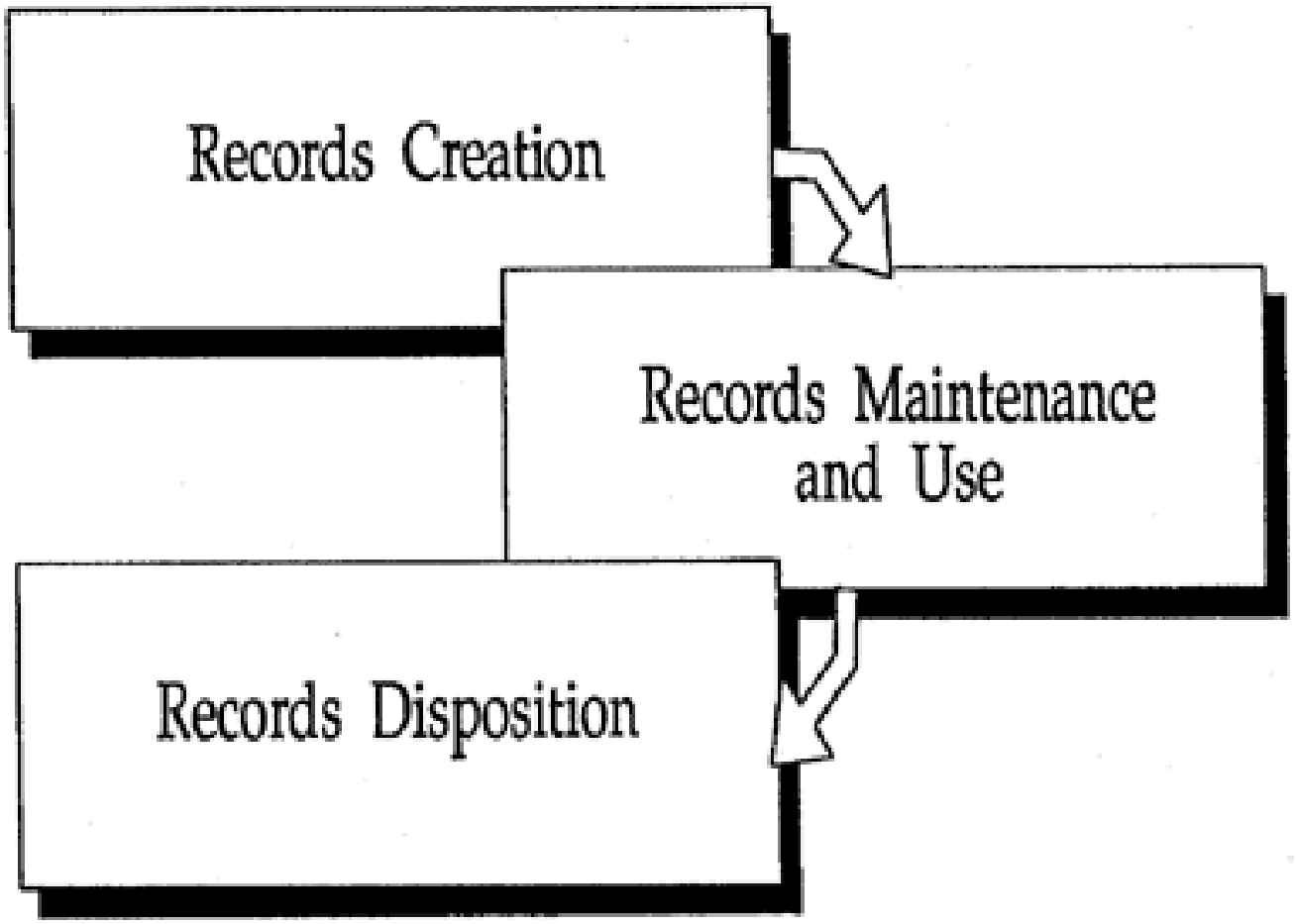

transistors
MOORE'S LAW
Dual-Core Intel® Itanium® 2 Processor
Intel® Itanium® 2 Processor
Intel® Itanium® Processor
Intel® Pentium® 4 Processor
Intel® Pentium® III Processor
Intel® Pentium® II Processor
Intel® Pentium® Processor
Intel486™ Processor
Intel386™ Processor
286
8086
8080
8008
4004
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
1970
1975
1980
1985
1990
1995
2000
2005
2010

The Functional View

- **Have it**
  - Identify the existence of information resources
  - Systematically assemble a collection
- **Find it**
  - Identify the works contained in the collection
  - Describe the content of the collection
  - Support intellectual access
- **Serve it**
  - Provide physical access

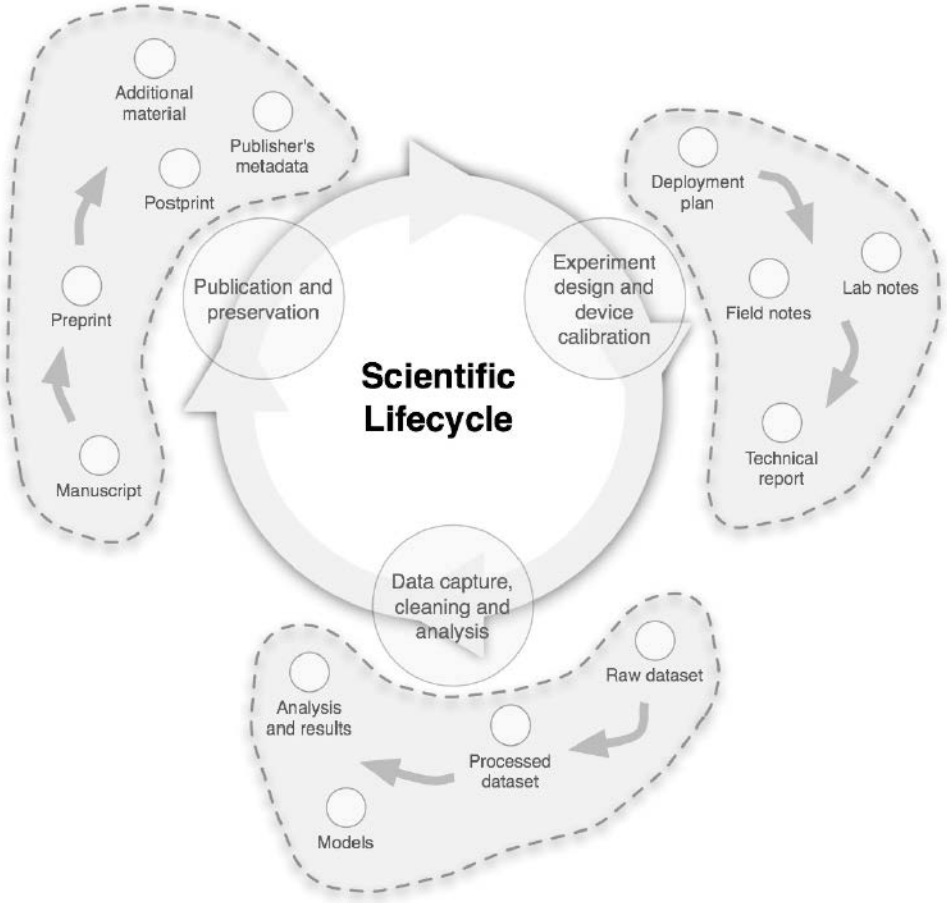
Hagler (1997), as interpreted in Taylor and Joudrey (2009)

NARA Records Life Cycle



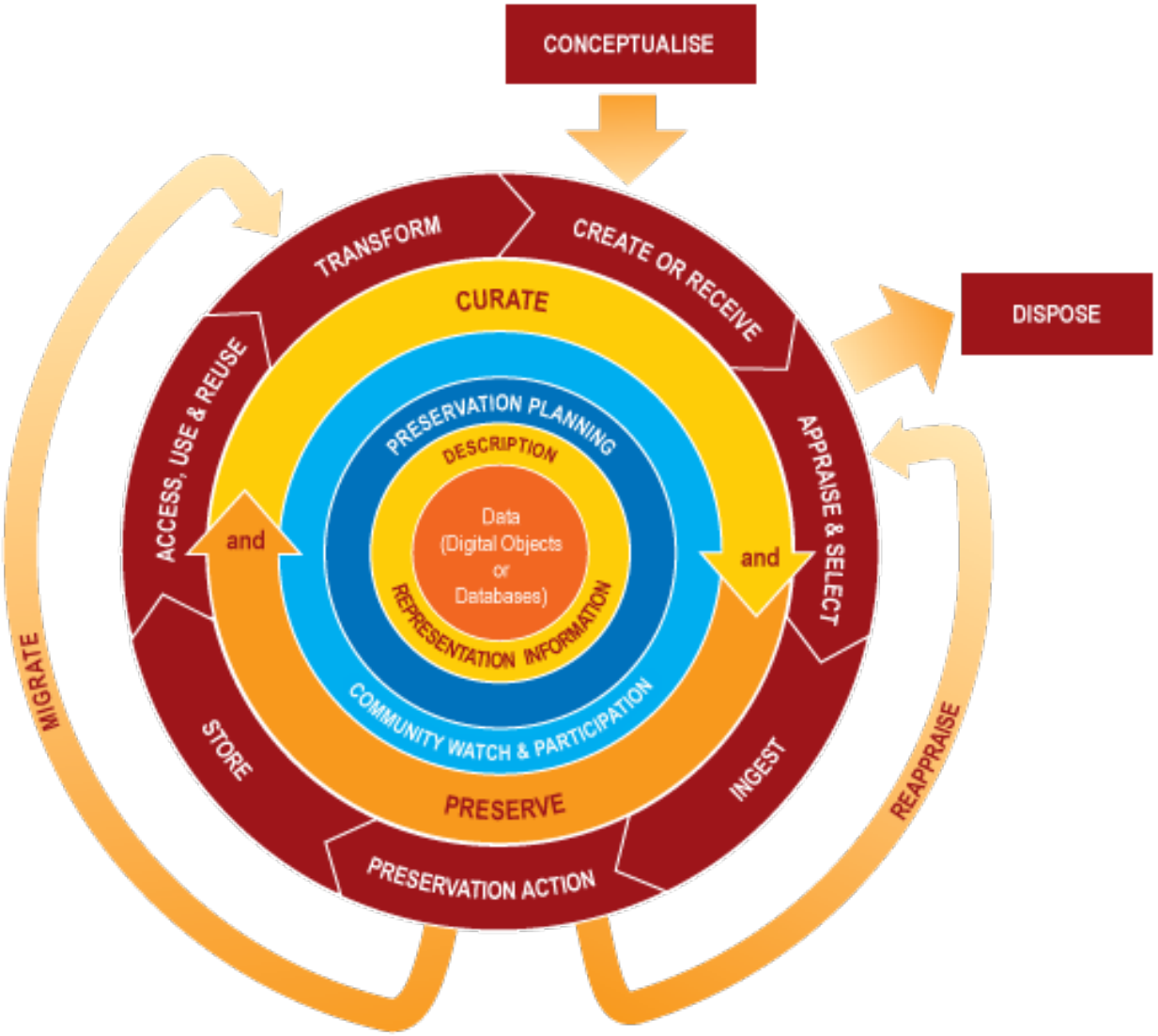
National Archives and Records Administration (2000)

A Scientific Information Lifecycle

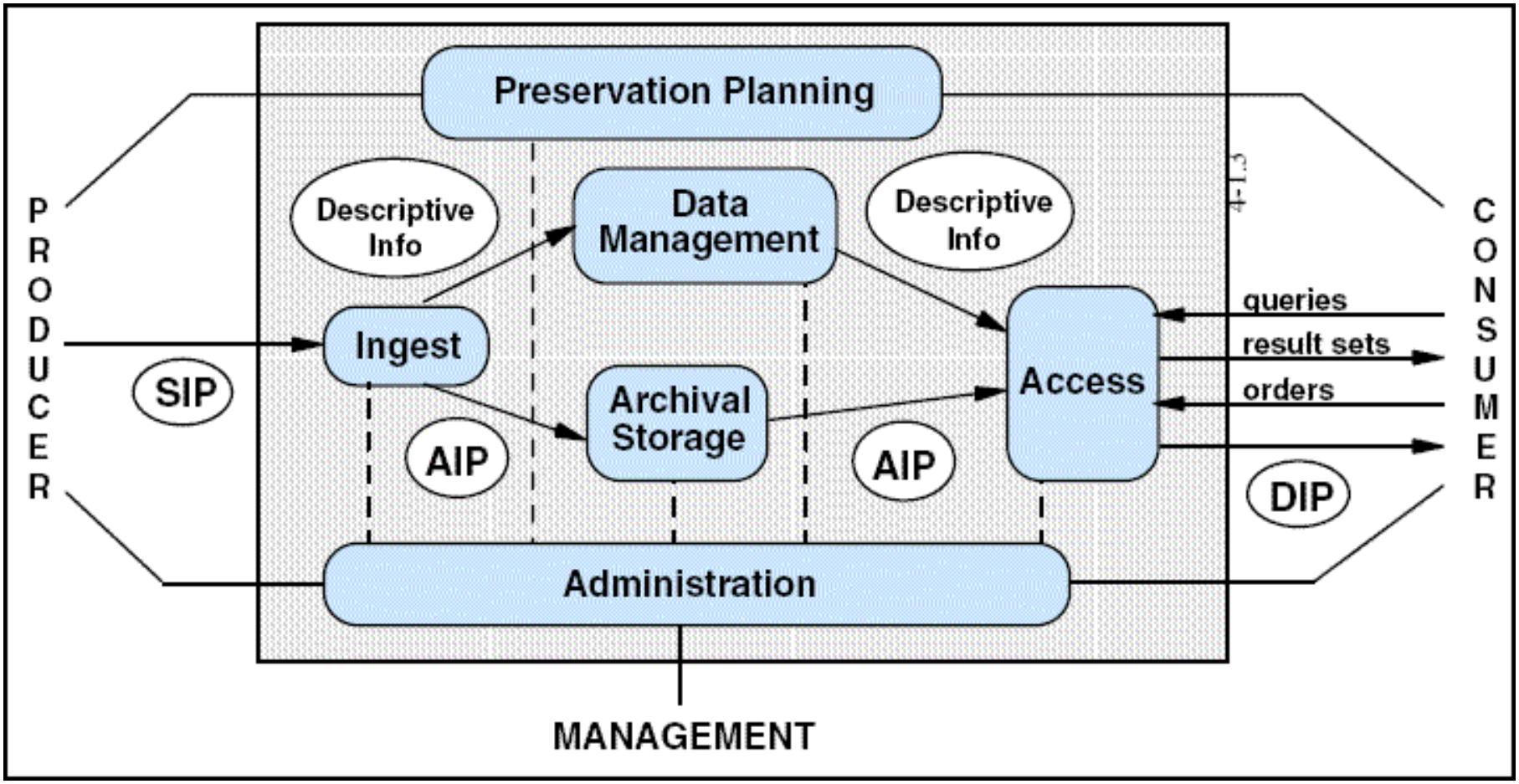


Alberto Pepe, AAHEP4 Summit (2010)

DCC Digital Curation Life Cycle



OAIS Reference Model



Types of “Metadata”

- Descriptive
  - Content, creation process, relationships
- Technical
  - Format, system requirements
- Usage
  - Display, derivative works
- Administrative
  - Acquisition, authentication, access rights
- Preservation
  - Media migration

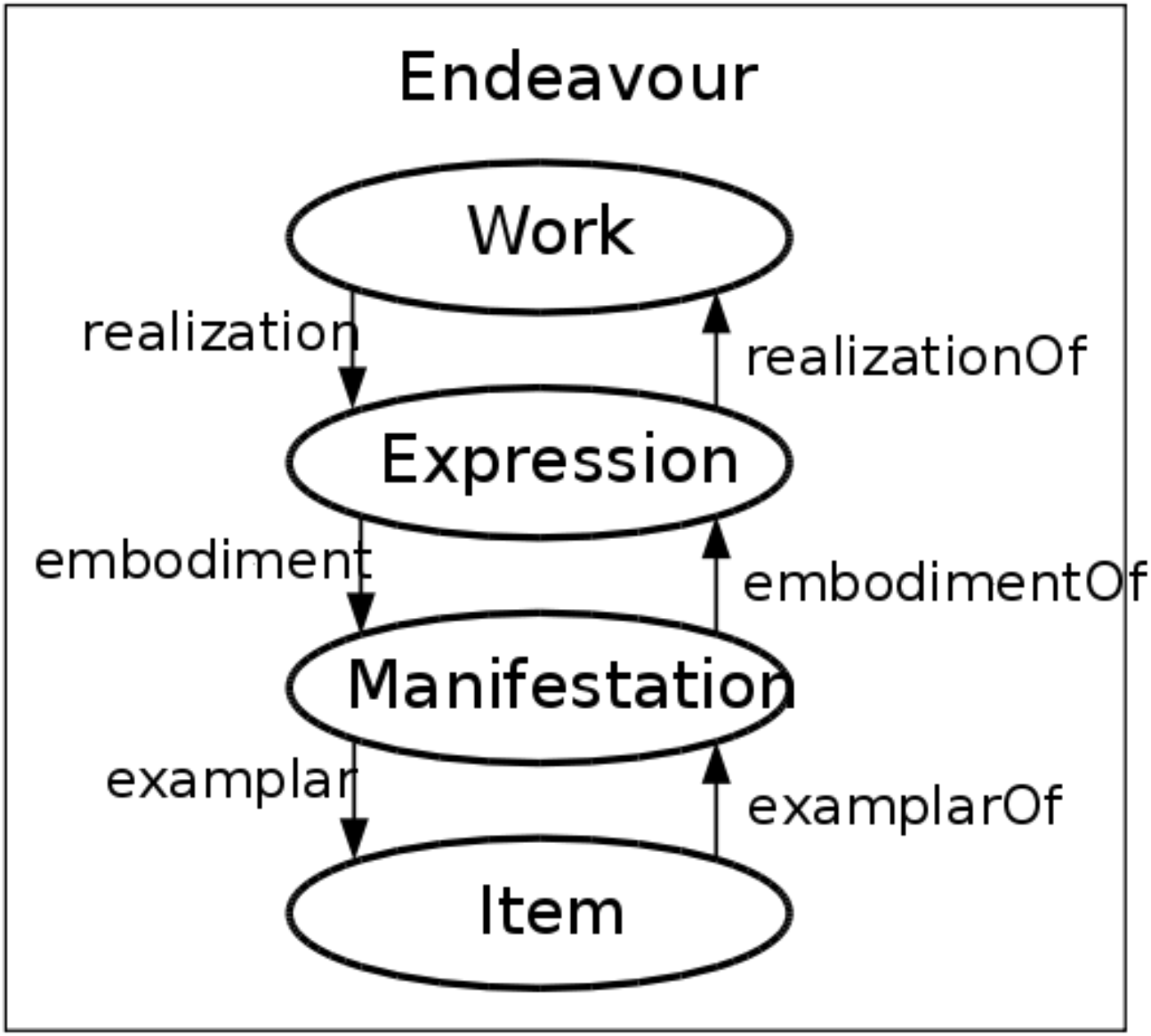
Adapted from Introduction to Metadata, Getty Information Institute (2000)

Five Levels of Metadata

- Framework
  - Functional Requirements for Bibliographic Records (FRBR)
- Schema
  - Dublin Core
- Vocabulary
  - Library of Congress Subject Headings (LCSH)
- Representation
  - Resource Description Framework (RDF)
- Serialization
  - RDF in eXtensible Markup Language (RDF/XML)

Adapted from Dante Alighieri, *Comedia* (c. 1321)

FRBR



The Organization of Information

Third Edition

Paperback
(ISBN 978-1-59…)

Copy 2
(barcode 102343…)

Dublin Core

- Title
- Creator
- Date
- Type
- Subject
- Language
- Identifier
- …

LCSH

INFORMATION FOR: Library science.

Narrower Term: Classification--Books--Library science
Narrower Term: Collectanea files
Narrower Term: Communication in library science.
Narrower Term: Comparative librarianship.
Narrower Term: International librarianship.
Narrower Term: Medical librarianship
Narrower Term: Mentoring in library science.
Narrower Term: Minorities in library science
Narrower Term: Music librarianship.

…

See Also: Bibliography.
See Also: Documentation.
See Also: Information science.

RDF XML

```
<?xml version="1.0"?>
< rdf:RDF
      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:dc= "http://purl.org/dc/elements/1.1/">

< rdf:Description rdf:about="http://www.w3schools.com">
   <dc:description>W3Schools - Free tutorials</dc:description>
   <dc:publisher>Refsnes Data as</dc:publisher>
   <dc:date>2008-09-01</dc:date>
   <dc:type>Web Development</dc:type>
   <dc:format>text/html</dc:format>
   <dc:language>en</dc:language>
< /rdf:Description>

< /rdf:RDF>
```

Some Challenges

- Evolution of information production
- Impersistence of access to digital content
  - Location, content, format, status
- Separation of content and services
- Invisibility to stakeholders
- Institutional boundaries (e.g., LAM)

Christine Borgman, The Invisible Library (2001)

My Homework P2

- Setting: Mission reconstruction app
- Collection: Apollo 15
  - Pre-flight planning
  - In-flight activities
  - Post-flight analysis
  - Post-flight recollections

Tonight

- What's this class about?
- Pieces of the puzzle
- ➢ All the usual stuff (syllabus, grading, …)

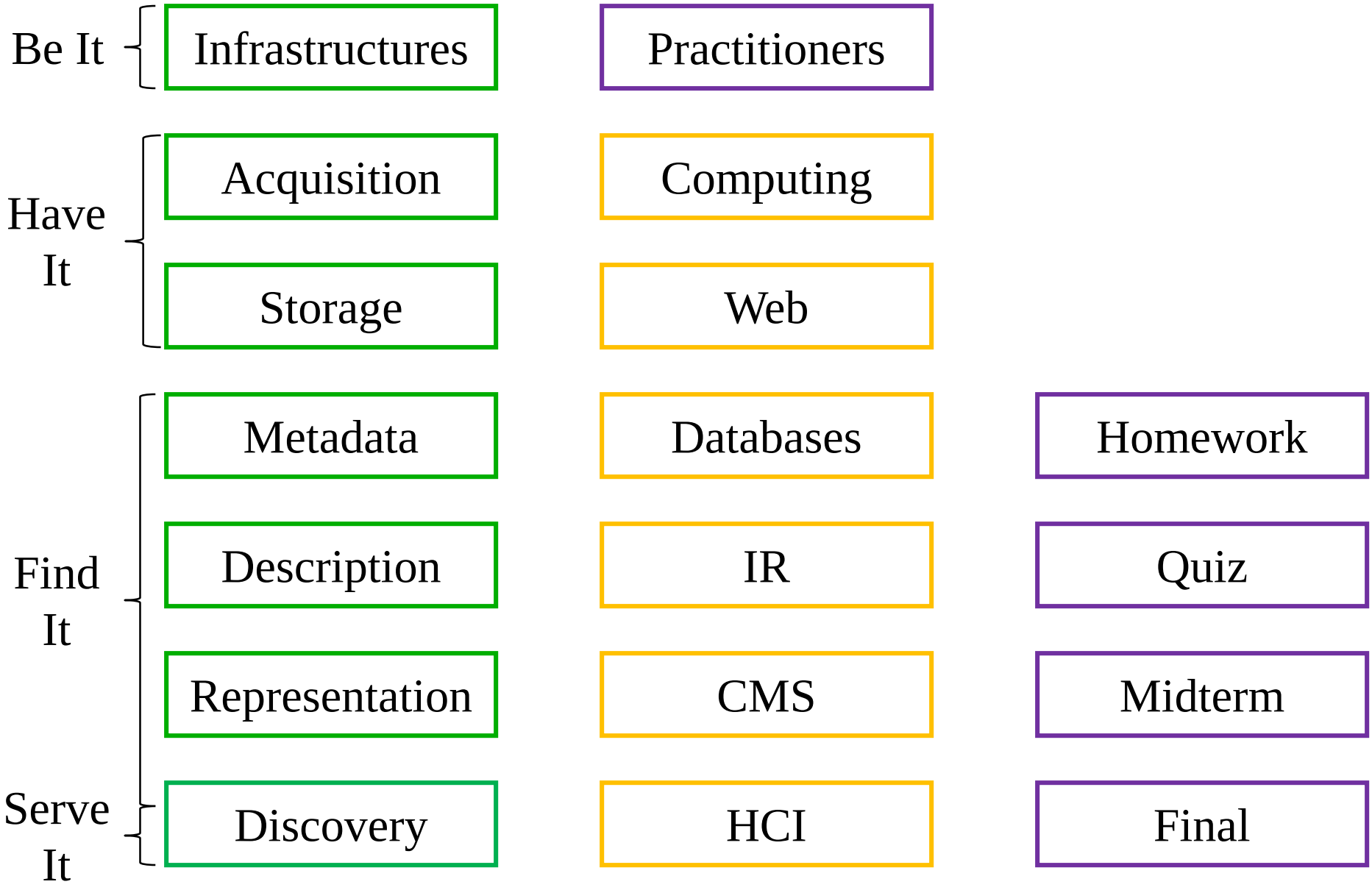
Class Structure

- Start promptly at 5:30, end promptly at 8:15
  - Except Dec 2 (**5:00-7:45 that day**)
- Typically, two breaks
  - 10 minutes after the first hour
  - 5 minutes after the second hour
- Bring a computer and use it
  - But stay focused on the class discussion
  - Current tuition+fees ≈ $133 per class session

Approach

- Readings (done before class)
  - Acquire background
  - Ready source for details
- Class sessions
  - Develop conceptual structure
- Homework (done after class)
  - Gain hands-on experience
- Quiz, exams
  - Focus effort, measure progress

The Grand Plan



Reading Strategies

- Assignments are found in two places:
  - Textbook chapters listed on schedule
  - Additional readings on a separate page
- Set aside an hour per assigned reading
  - Not all on the same day!
- Read initially for understanding, not detail
  - You can find details later (if you know where!)

Grading

- 50%-55% individual work
  - Exams: 35% for the best, 15% for the other
- 45%-50% your choice (individual or group)
  - 5% each for best 10 of the 11 homework/quiz
  - First and last homework are graded pass/fail
  - Others (and quiz) graded on a 0-5 scale
- No curve
  - 90-100 some sort of A, 80-89 some sort of B, …

The Fine Print

- Group work is encouraged on homework
  - But you must personally write what you turn in
- Deadlines are firm and sharp
  - Allowances for individual circumstances are included in the grading computation
- Academic integrity is a serious matter
  - No group work during the exams or the quiz!
  - Scrupulously respect time limits

Finding Me

- Doug Oard
- Office: HBK 2118F
  - I'm normally there from 4:45-5:15 on class days
  - I'm also normally happy to stay after class
  - People in the E-Discovery lab know if I am around
- Email: oard@umd.edu
  - Expect an answer within 24 hours
  - Include a phone number if discussion would help
  - Indicate if you do not want answers sent to the class

A Personal Approach to Learning

- Work ahead, so that you are never behind
- Find new questions everywhere
  - Then find the answers somewhere
- Enrich your practical skills relentlessly
- Pick topics you want to learn more about

Before You Go

On a sheet of paper, answer the following (ungraded) question (no names, please):

What was the muddiest point in today's class?