



College of Information Studies

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

Information Infrastructures

Week 1

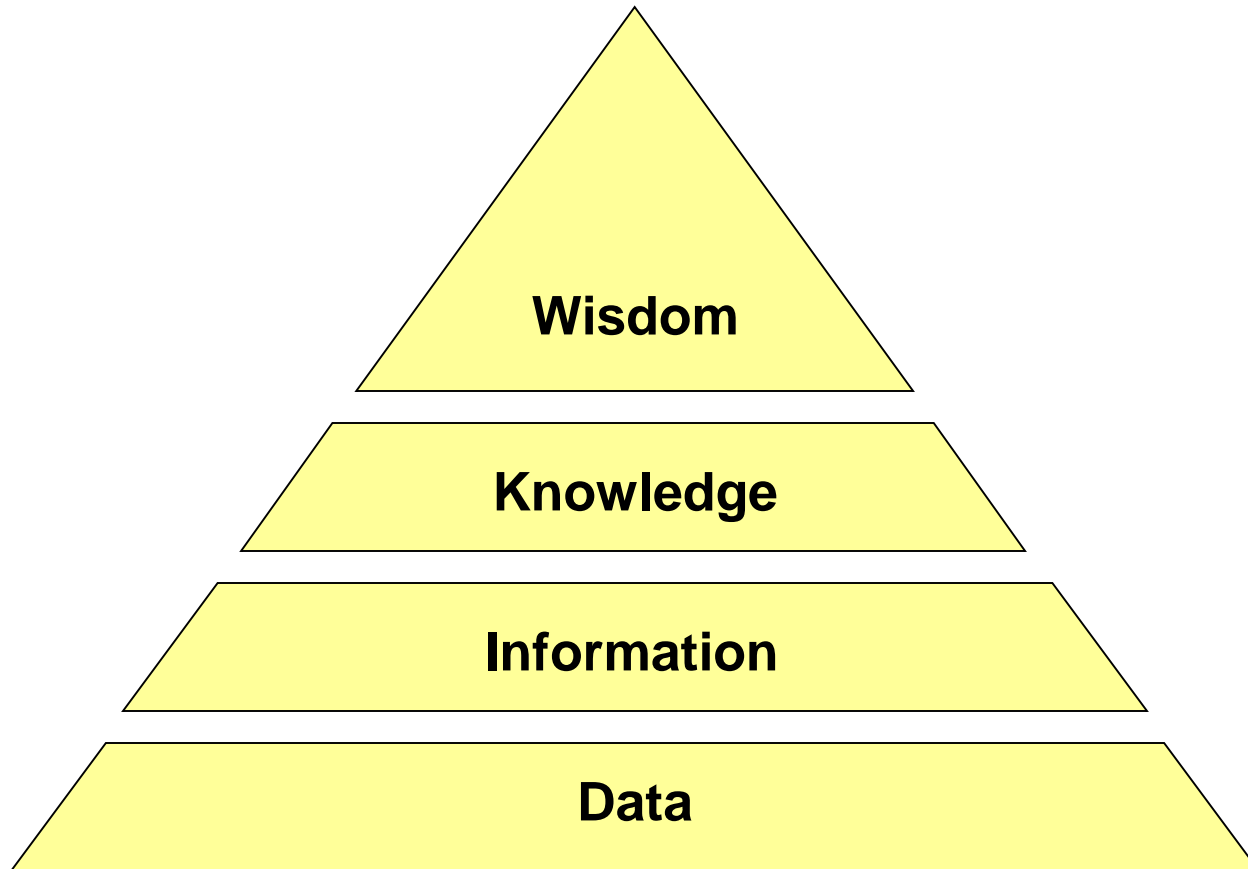
LBSC 671

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
 - (and visible when it fails!)



“Information Institutions”

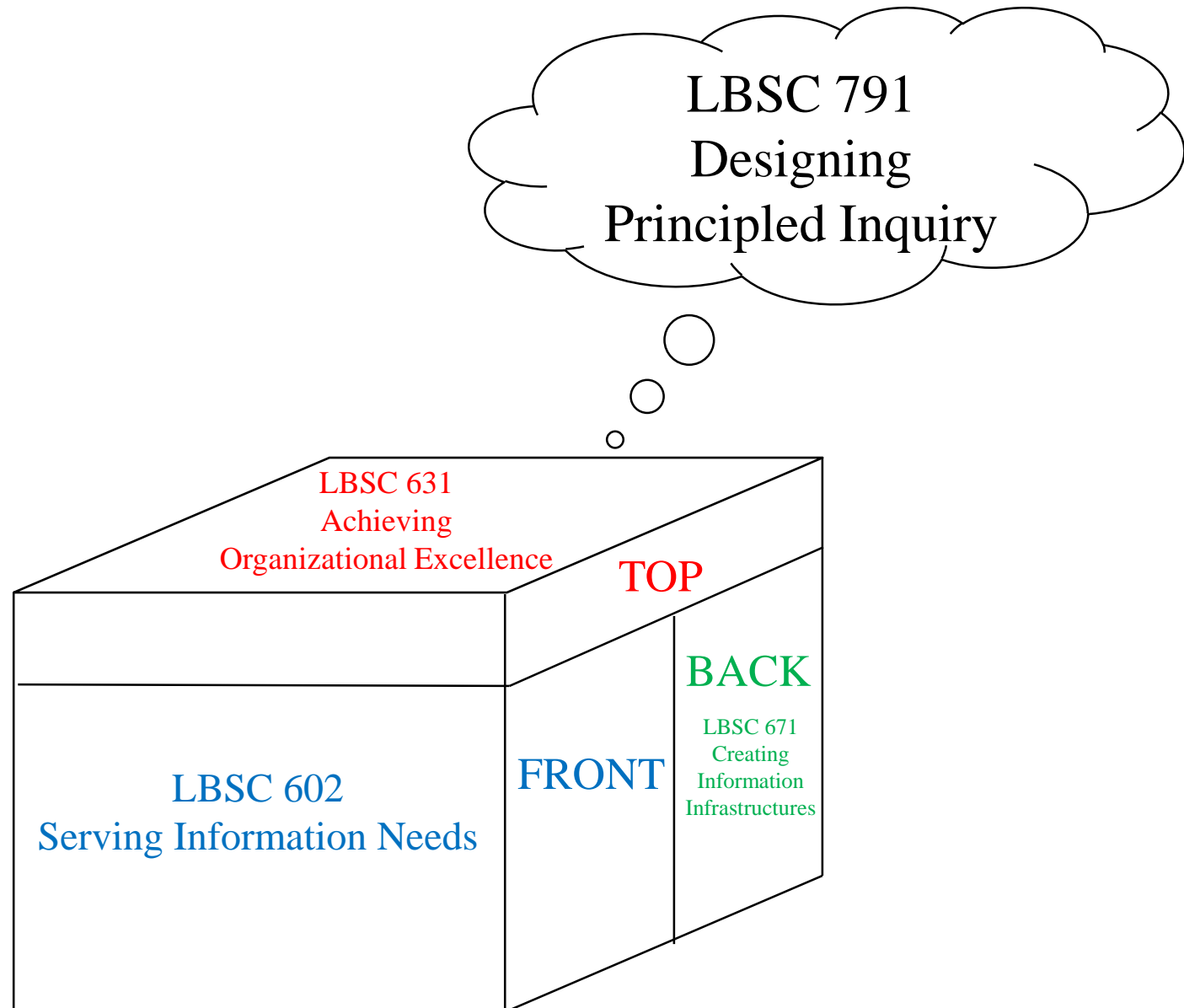
- Schools
 - Primary, Secondary, Tertiary
- Libraries
 - Public, Academic, School, Special
- Archives
 - Government, Corporate, ...
 - Special collections, Historical societies
- Museums
 - Art, Material culture, Natural history
- Search engines
 - Google, Bing, Baidu, Yandex

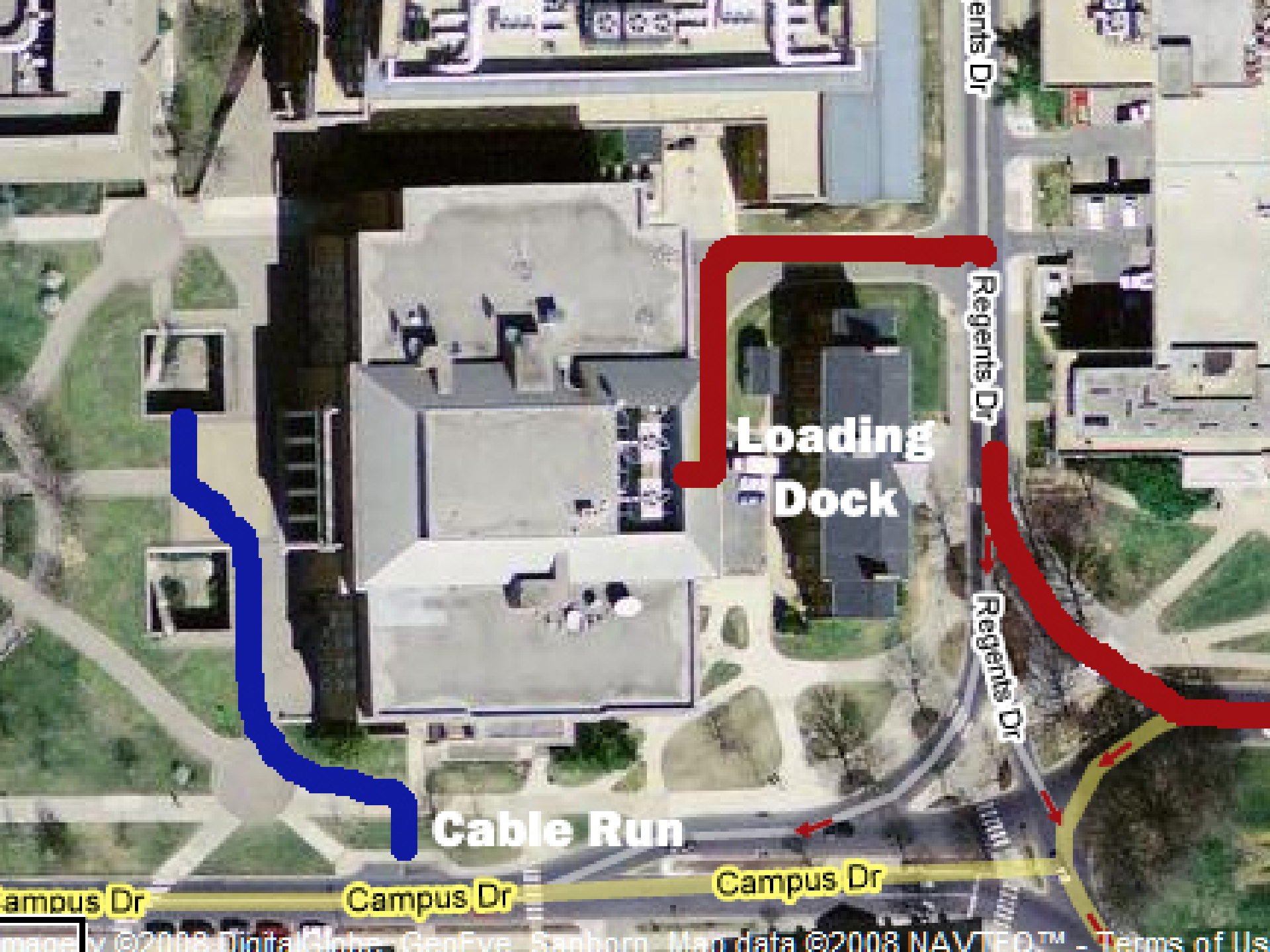
“LAM”

LBSC
671

LBSC
671

The Core





Loading Dock

Cable Run

Campus Dr

Campus Dr

Campus Dr

Regents Dr

Regents Dr

Regents Dr

The MLS Program

- Specializations

- Top {
 - School Library
 - Archives, Records, and Information Management
 - History and Library Science dual degree program
 - Government Information Management & Services
- Front {
 - Information and Diverse Populations
- Back {
 - Curation and Management of Digital Assets

- Unspecializations

- 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
- INFMM
 - 741: Social Computing
 - 743: Internet Applications

Tonight

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

THE WORLD'S CAPACITY TO STORE INFORMATION

This chart shows the world's growth in storage capacity for both analog data (books, newspapers, videotapes, etc.) and digital (CDs, DVDs, computer hard drives, smartphone drives, etc.)

In gigabytes or estimated equivalent

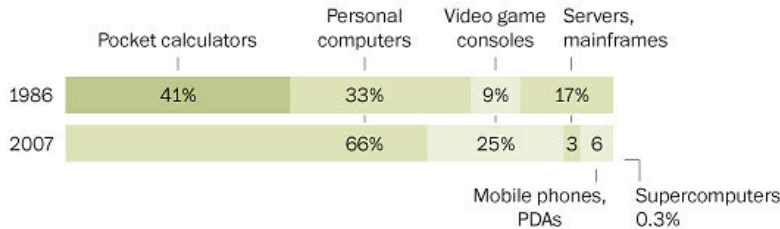
1986
ANALOG
2.62 billion

DIGITAL
0.02 billion

COMPUTING POWER

In 1986, pocket calculators accounted for much of the world's data-processing power.

Percentage of available processing power by device:



2007
ANALOG
18.86 billion gigabytes

Paper, film, audiotape and vinyl: 6.2%

Analog videotapes: 93.8%

Other digital media: 0.8%*
Portable media players, flash drives: 2%
Portable hard disks: 2.4%

CDs and minidisks: 6.8%

Computer servers and mainframe hard disks: 8.9%

Digital tape: 11.8%

DVD/Blu-ray: 22.8%

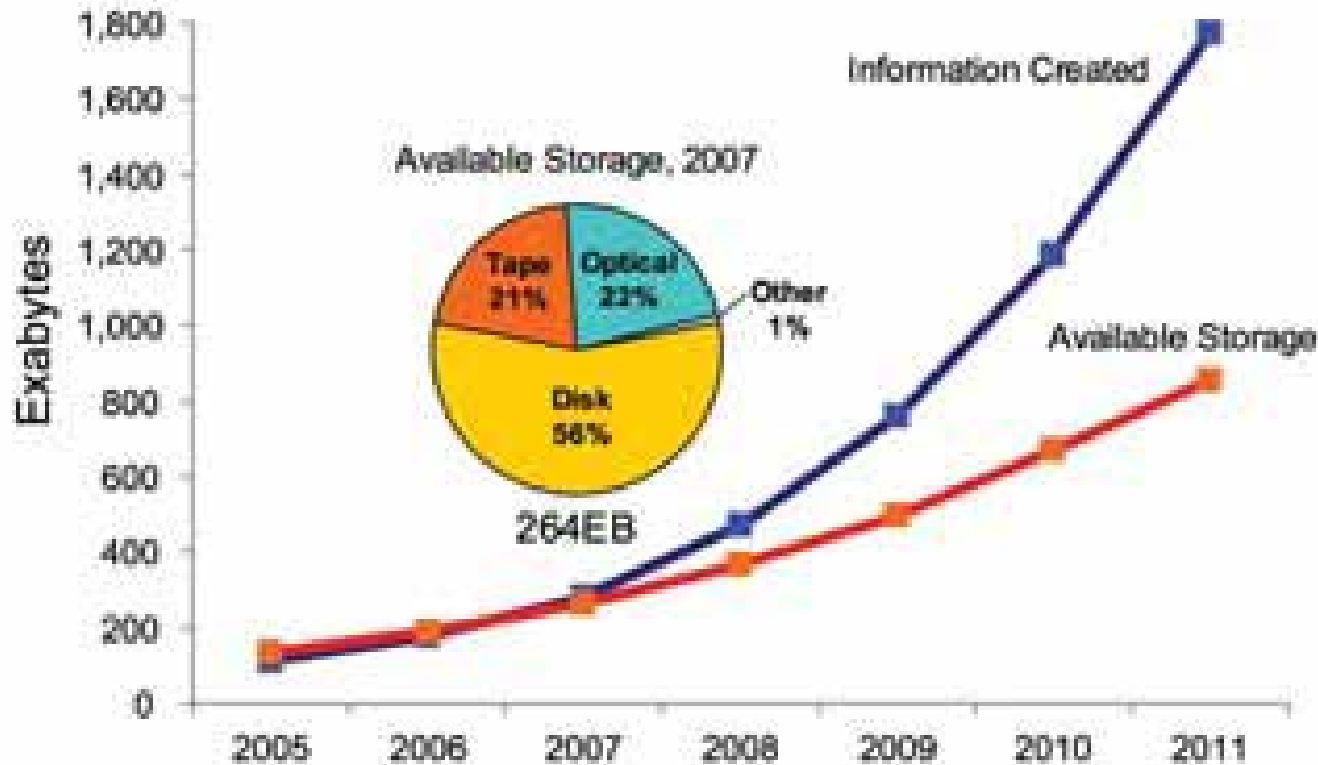
PC hard disks: 44.5%

123 billion gigabytes

*Other includes chip cards, memory cards, floppy disks, mobile phones/PDAs, cameras/camcorders, video games

2007
DIGITAL
276.12 billion gigabytes

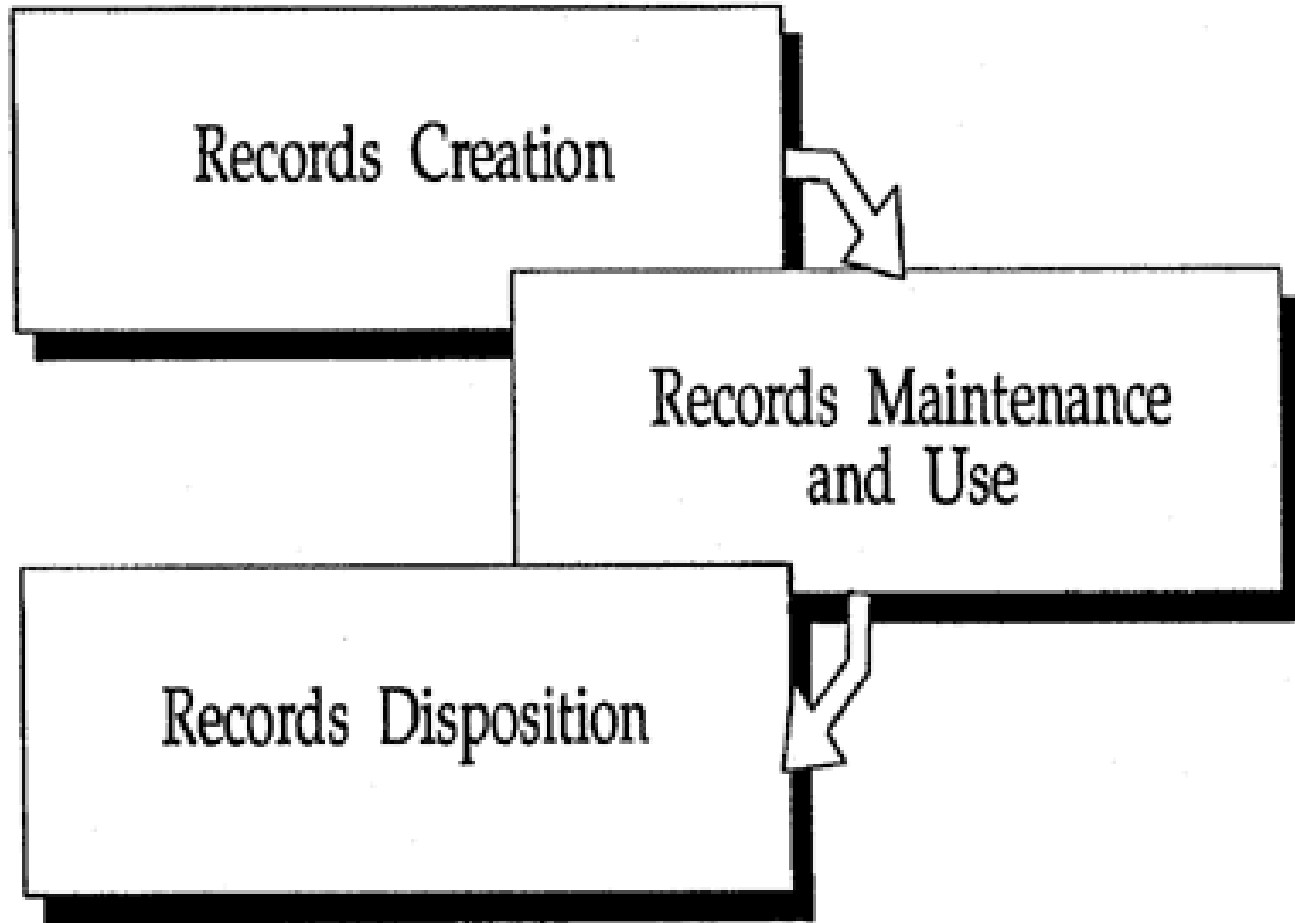
“Data In Motion” vs. “Data At Rest”



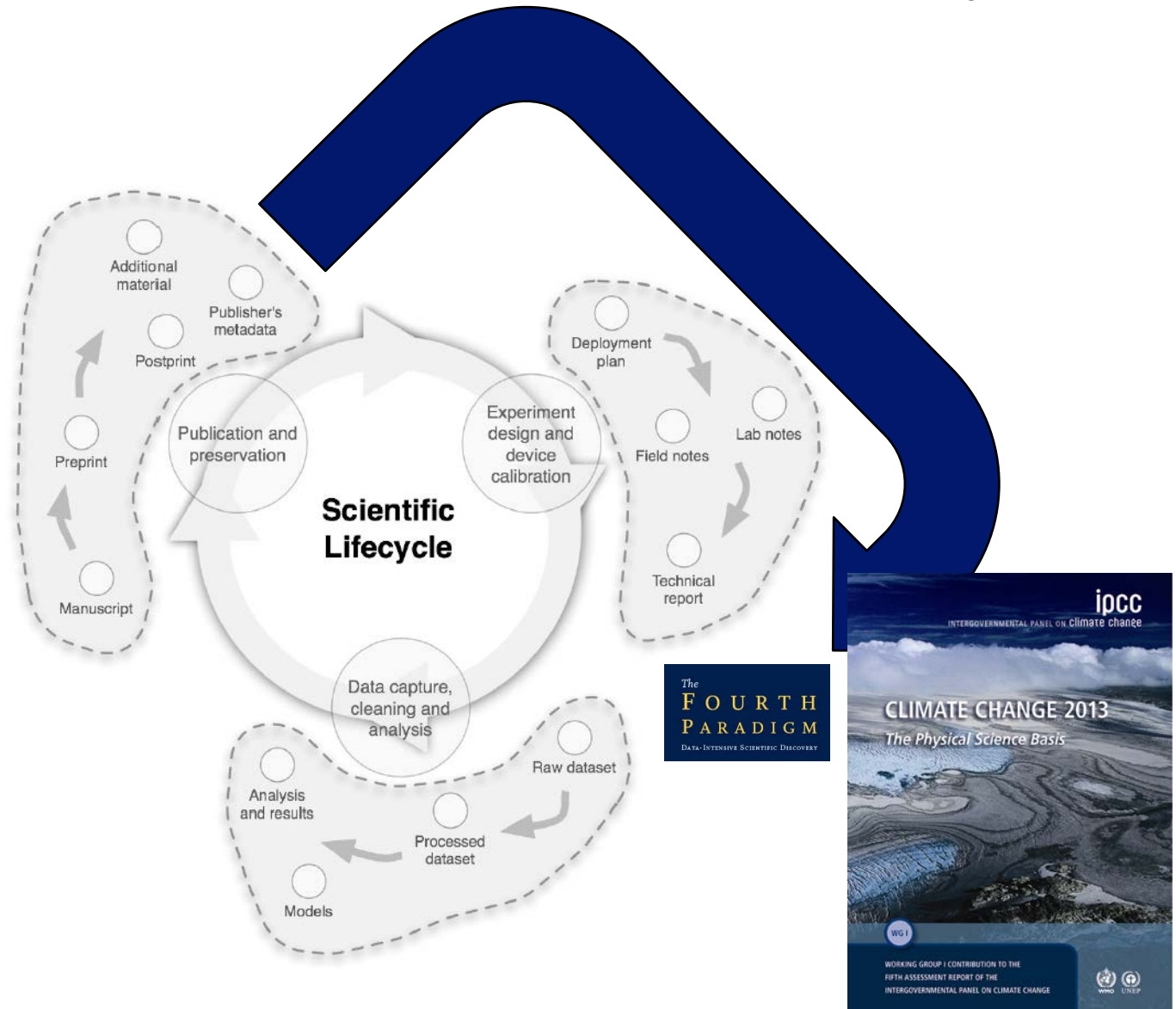
The Functional View

- **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

NARA Records Life Cycle

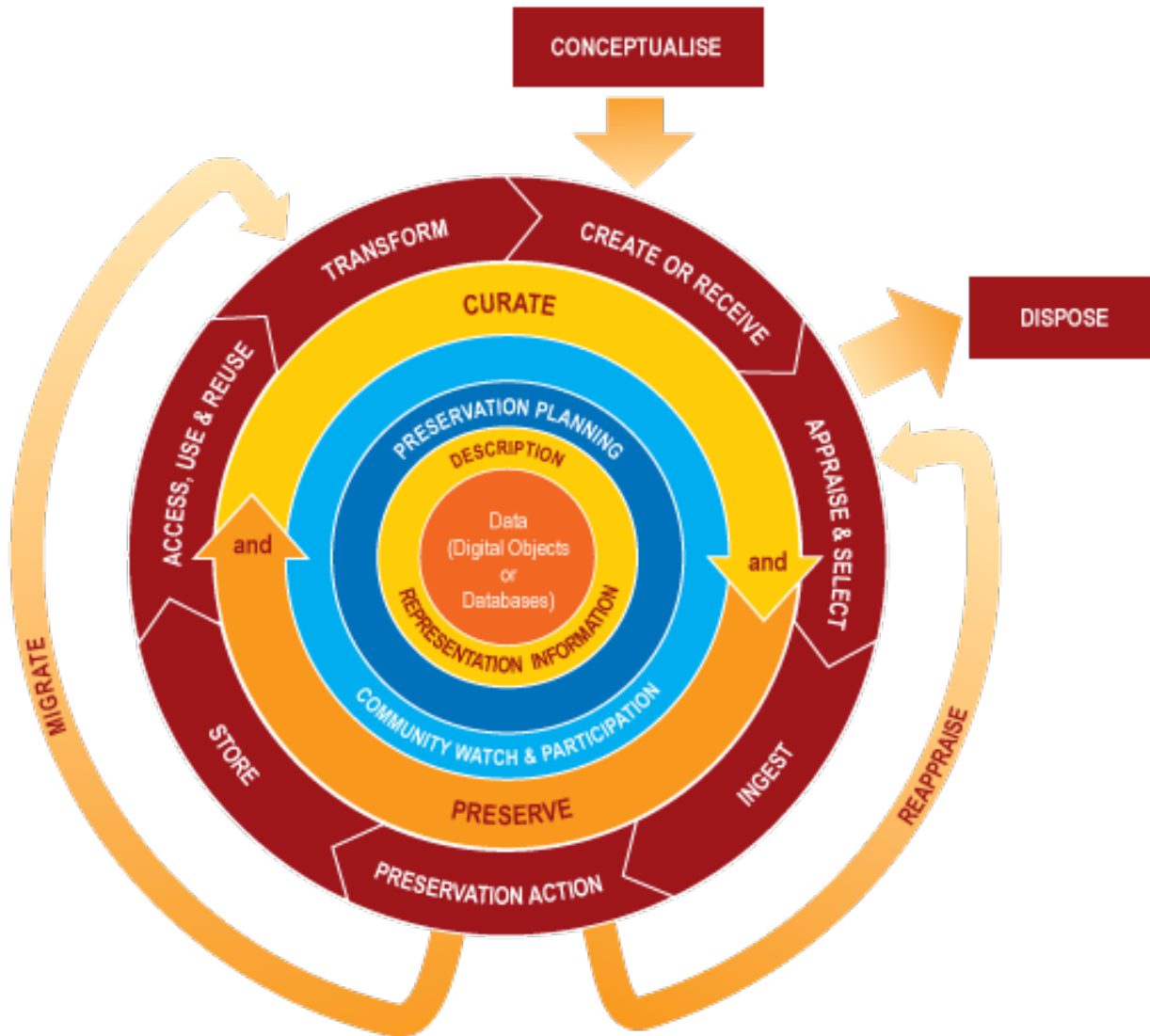


A Scientific Information Lifecycle

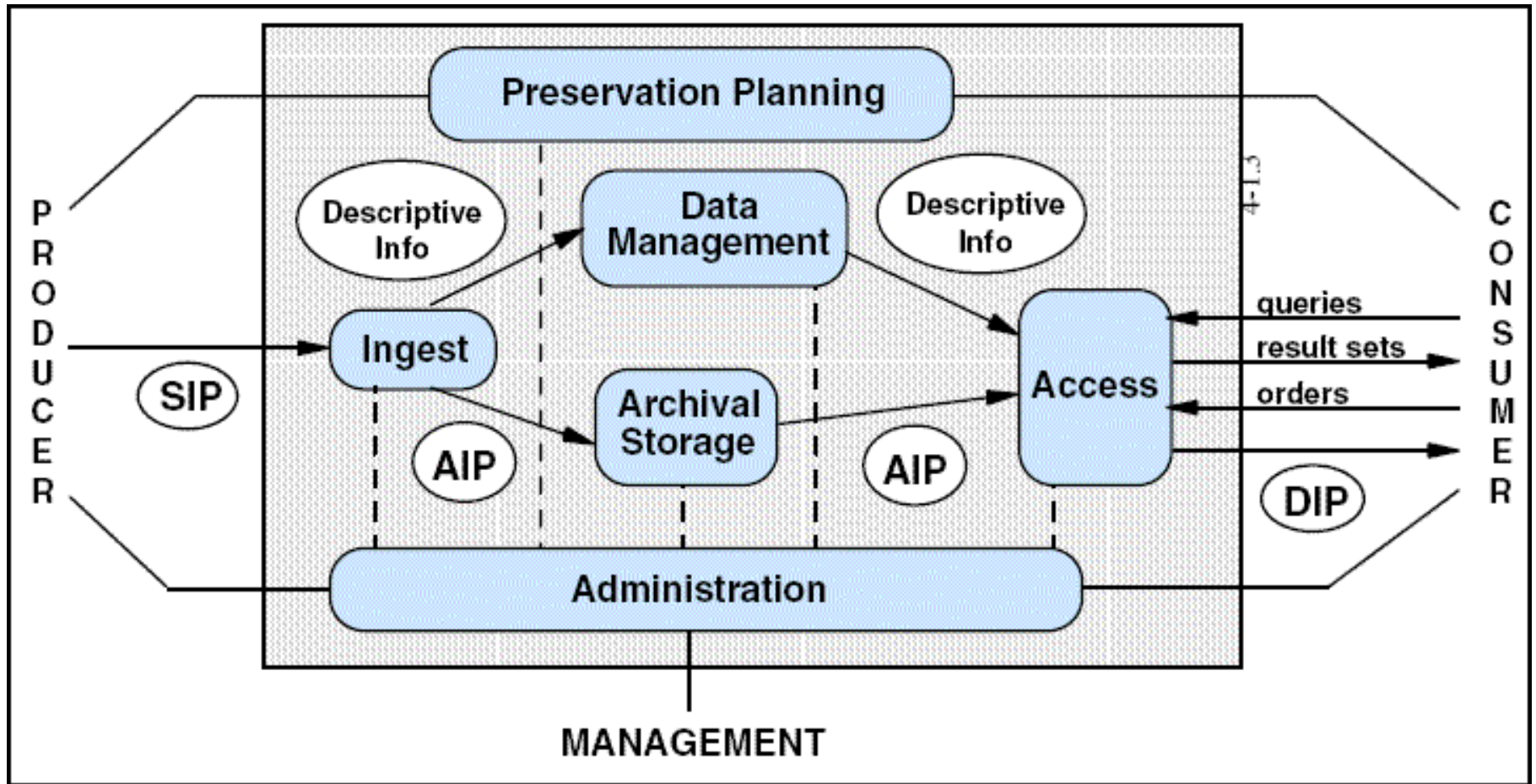


Alberto Pepe, AAHEP4 Summit (2010)

DCC Digital Curation Life Cycle



OAIS Reference Model



Some Types of “Metadata”






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

Not in
Taylor &
Joudrey

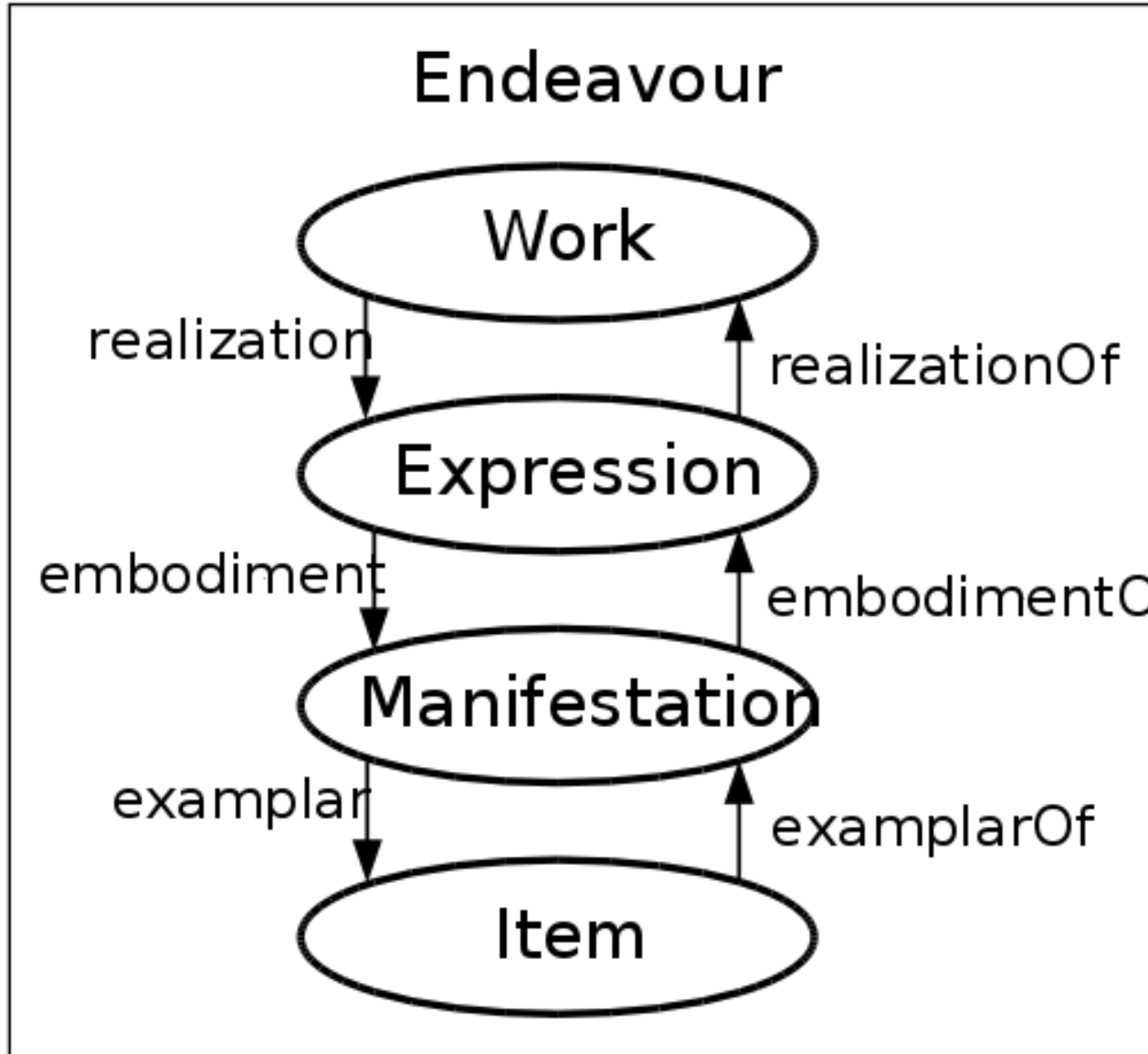


Adapted from
Introduction to Metadata,
Getty Information Institute (2000)

Aspects of Metadata

- Framework 
 - Functional Requirements for Bibliographic Records (FRBR)
- Schema (“Data Fields and Structure”) 
 - Dublin Core
- Guidelines (“Data Content and Values”) 
 - Resource Description and Access (RDA)
 - Library of Congress Subject Headings (LCSH)
- Representation (abstract “Data Format”) 
 - Resource Description Framework (RDF)
- Serialization (“Data Format”) 
 - RDF in eXtensible Markup Language (RDF/XML)

A Framework: FRBR



The Organization
of Information

Third Edition

Paperback
(ISBN 978-1-59...)

Copy 2
(barcode 102343...)

A Schema: Dublin Core



Content

- Title
- Subject
- Description
- Type
- Coverage [spatial, temporal, ...]
- Related resource
- Rights

Instantiation

- Date [Created, Modified, Copyright, ...]
- Format
- Language
- Identifier [URI, Citation, ...]

Responsibility

- Creator
- Contributor
- Source
- Publisher

A Guideline: 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.](#)

A Serialized Representation: 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>
```

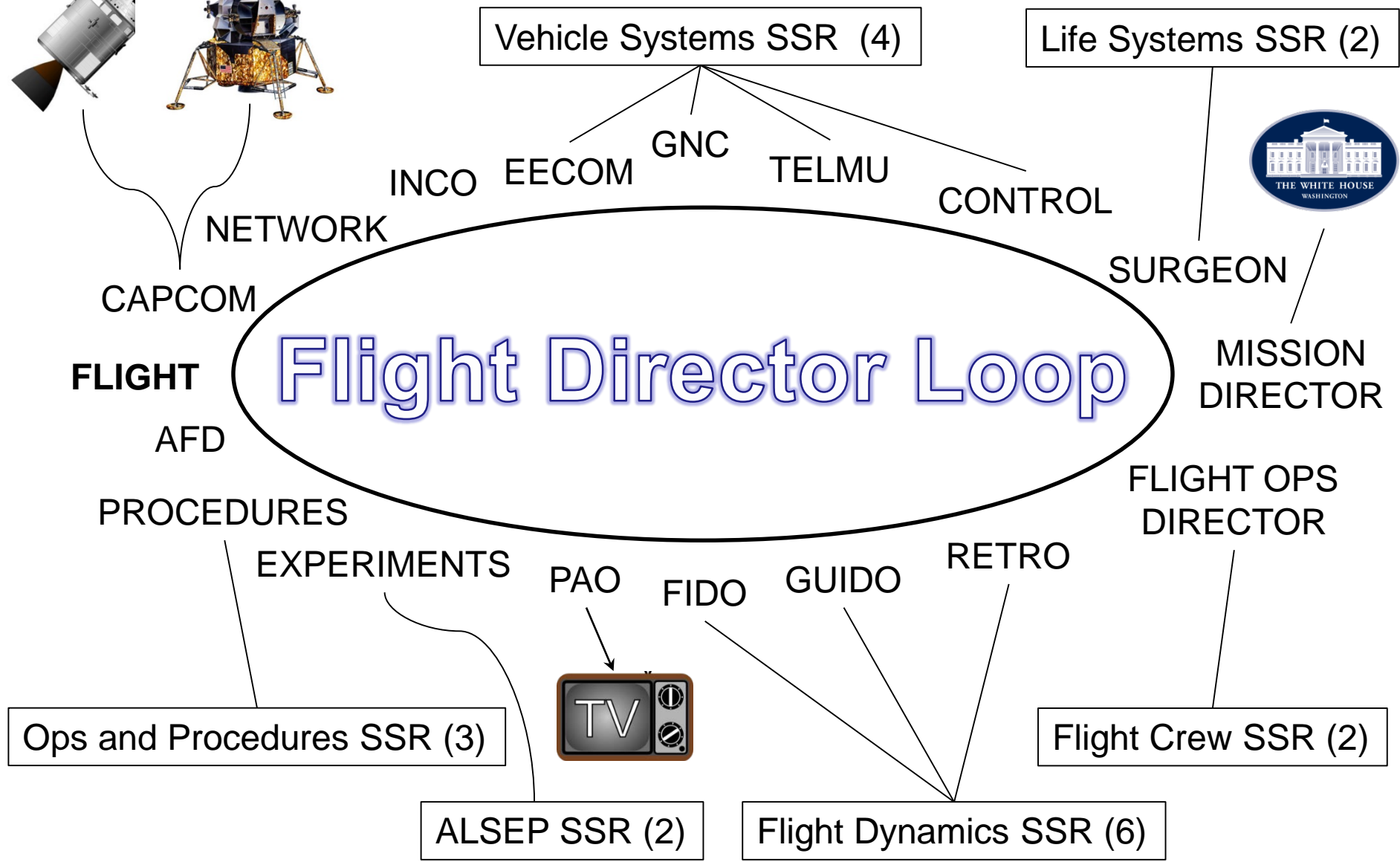
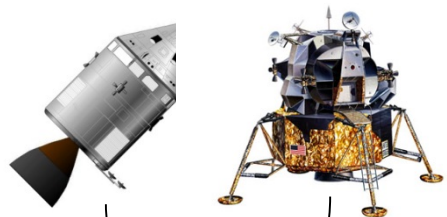
Some Challenges

- 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)

My Homework P2

- Setting:
 - Mission reconstruction app
- Collection:
 - Recorded audio from the Apollo Program
 - Mission control discussions
 - Radio communications
 - Onboard the spacecraft





Flight Director Loop





0A 1268

MULTI-CHANNEL REPRODUCER
SOUNDScriber
MODEL NO. 1000
SERIAL NO. 1000
MFG. BY THE REPRODUCER CO.
NORTH HAVEN, CONN. U.S.A.

TRIG A OUT

618

AMPEX

AMPEX

AMPEX



STANLEY
CORP. MANHATTAN

Apollo 7 Track Sheet

APOLLO 7 AS-205

HISTORICAL RECORDER = 1

10-08-66

CH 1	TIME GMT IRIG B FORMAT		CH 16	REAL TIME CMD CONTROL	POS 816
2	RTC LOOP	L-113	17	TLM INSTA CONTROL	POS 819
3	GOSS 4 LOOP	I-055	18	PAO RELEASE LOOP	I-299
4	NETWORK COORD LOOP	L-107	19	RTCC CMD NTWK LOOP	L-025
5	SPACE ENVIRONMENT	POS 90	20	MOCR SYS 2 LOOP	L-022
6	FLT OPS DIR	POS 720	21	NETWORK CALL LOOP	L-033
7	DOD PIO	POS 79	22	NASA RECOVERY COORD	POS 82
8	COMM CONTROLLER	POS 205	23	CMD LOAD CONTROL	POS 815
9	SPACEFLIGHT METEOROLOGY	POS 91	24	TLM INST CONTROL	POS 818
10	ASST NASA RCYVY COORD	POS 83	25	COMMUNICATIONS TECHNICIAN	POS 820
11	RECOVERY DISPLAY CONTROL	POS 84	26	NAM SAM ASST	POS 76
12	GOSS 3	I050	27	MEAFSA	POS 77
13	RECOVERY DISPLAY CONT	POS 641	28	GOSS 9	1098
14	DOD ASST FOR COMM-2	POS 89	29	GOSS 11	1066
15	CMD LOAD CONTROL	POS 817	30	VOICE ANNOTATION	

Tonight

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

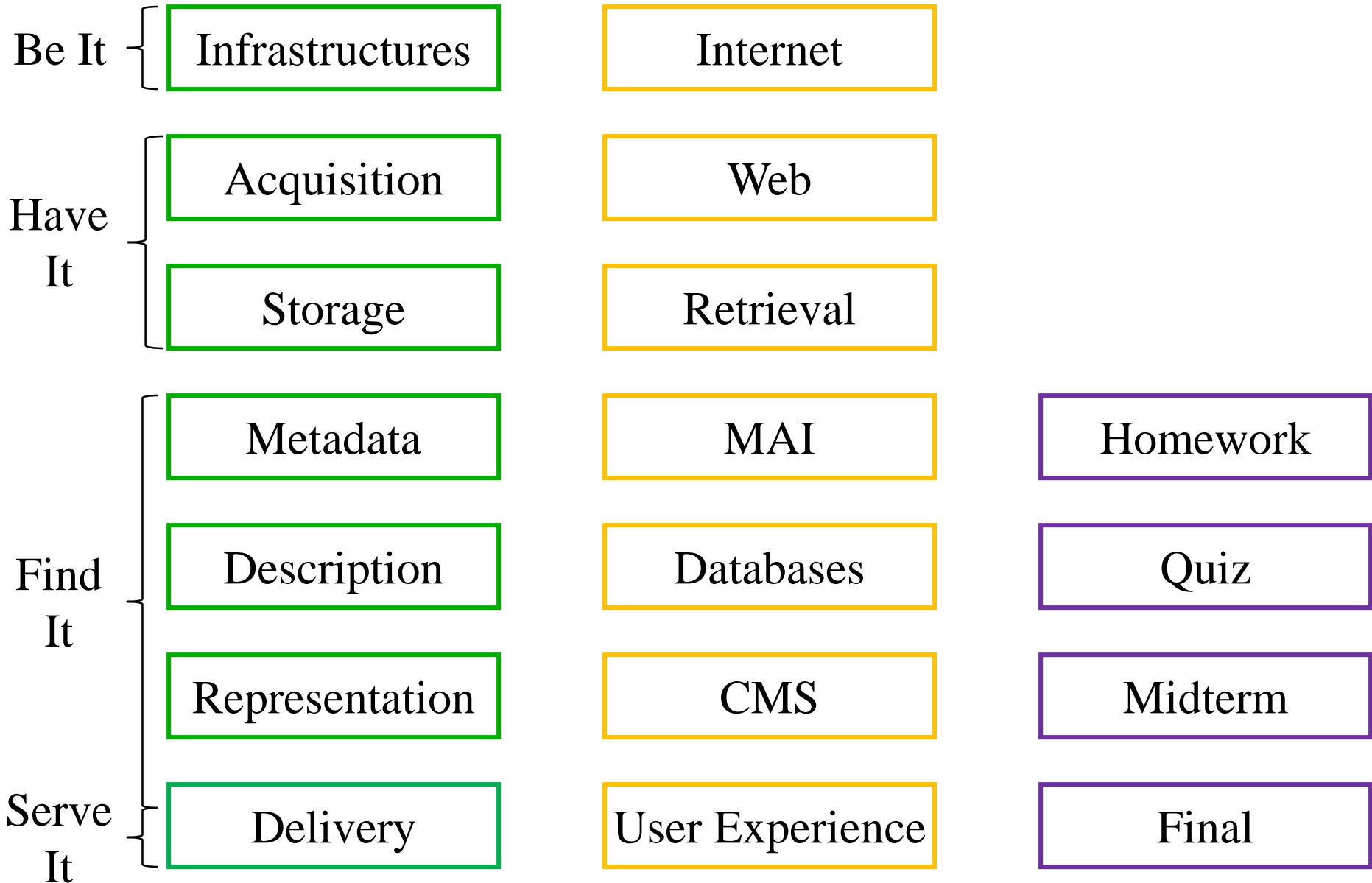
Class Structure

- Start promptly at 5:30, end promptly at 8:15
 - Today and on Feb 3 we may end a bit early
- 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 \approx \$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

- Readings are linked from the schedule
 - Typically 3 per week
- Set aside an hour per reading
 - Not all on the same day!
- Read initially for understanding, not detail
 - You can find details later (once you know where!)

Grading

- 52%-56% individual work
 - Exams: 35% for the best, 15% for the other
 - Graded on a 100-point scale
- 44%-48% your choice (individual or group)
 - 4% each for best 12 of the 13 homework/quiz
 - First and last homework are graded pass/fail
 - Others (and quiz) graded on a 0-4 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

- Office: HBK 2118F
 - I'm normally there from 4:45-5:15 on class days
 - I'm also usually happy to stay after class
 - People in the lab across from my office know when 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 my answer 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?