



Preservation Models & Packages

Developed in partnership with the



Digital Preservation Coalition

This POWRR Institute is generously funded by the



INSTITUTE of
Museum and Library
SERVICES

Expected Outcomes

Preservation Models & Packages

- ✓ Explain at a high level the main components of the OAIS standard
- ✓ Describe the elements of the information model and their relevance to the preservation lifecycle
- ✓ Design a basic information package and select relevant metadata standards

Why Do We Need Models?



www.digitalbevaring.dk

- High-level conceptual map for activities
- Can help set requirements
- Supports identification and development of standards
- Framework for comparing and assessing approaches

What Are the Risks?

- Media obsolescence
- Media failure or decay (such as “bit rot”)
- Natural / human-made disaster
- File format obsolescence



Images by Aldric Rodríguez Iborra, Erin Standley, Marie Van den Broeck, Edward Boatman and Dillon Choudhury from the Noun Project

Stuff Happens



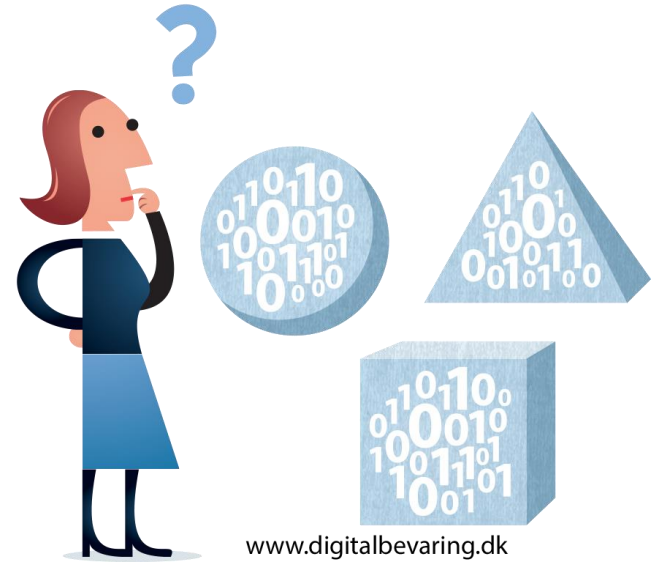
Whenever a digital collection is moved, processed, curated or altered in any way.... things can go wrong!

- Network dropouts at critical times
- Disks get full, subsequent data copied there is lost
- Software bugs lead to unexpected results
- Human error leads to all sorts of issues

Stuff happens a lot more at scale!

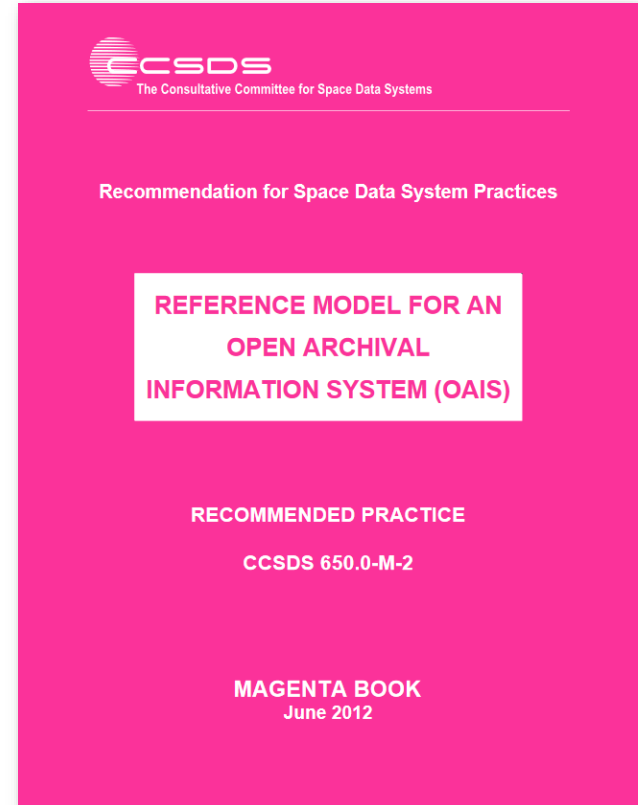
How Do We Solve These Problems?

- Keep more than one copy
- Refresh storage media
- Know what you have
- Integrity check your data (also called “Fixity”)
- Use ‘open’ formats
- Carry out preservation actions



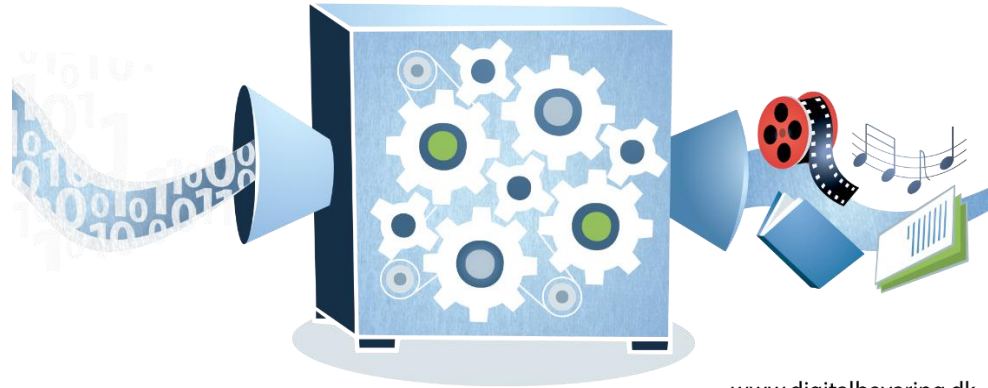
What is OAIS?

- Open Archival Information System Reference Model
- Originally developed by Consultative Committee for Space Data Systems
- An international standard
ISO 14721:2012
- Vocabulary and basic framework for much digital preservation work

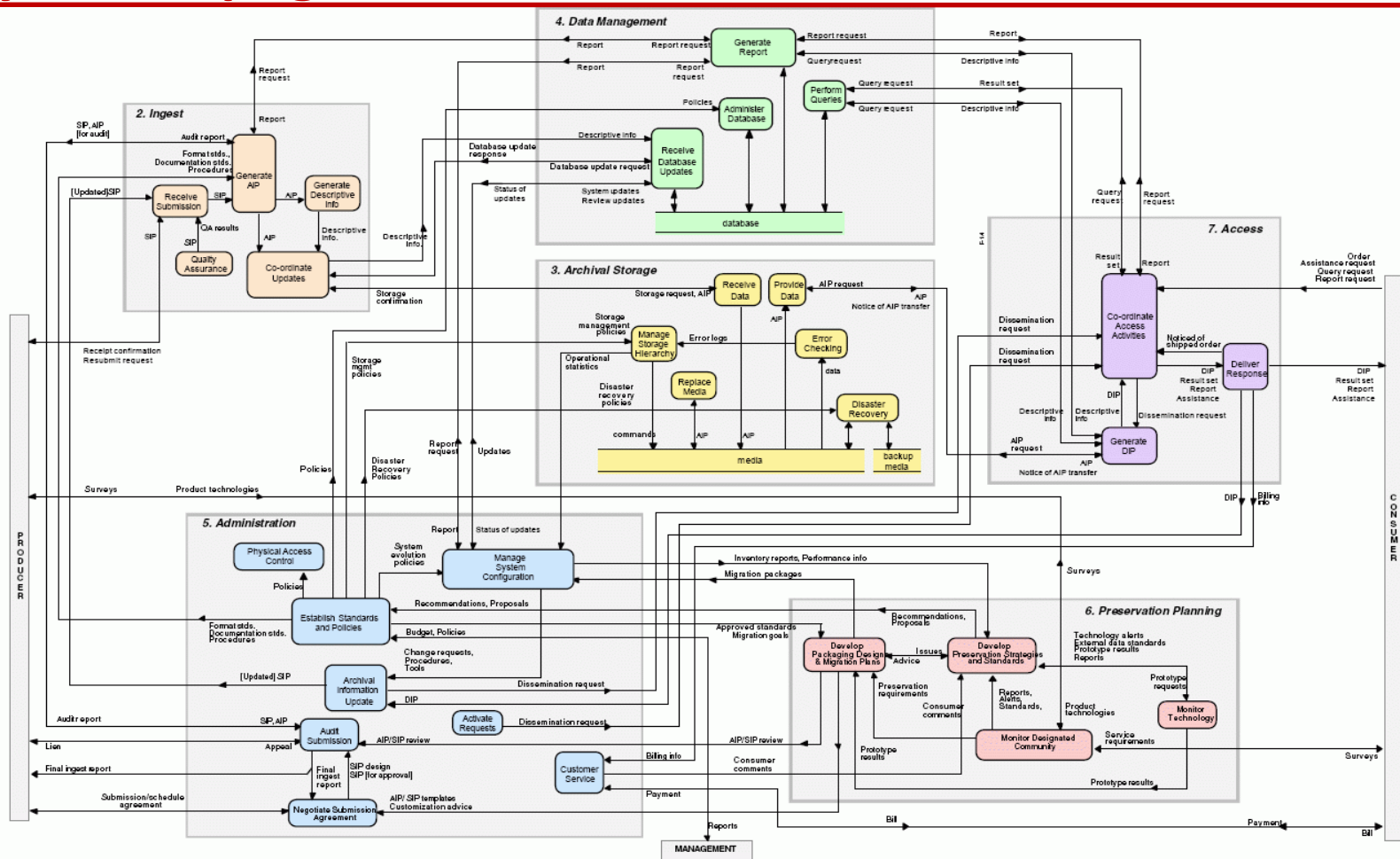


Basic Definition of an OAIS

a **reference model** ... to establish a **system** for archiving information, both digitalized and physical, with an **organizational scheme** composed of people who accept the **responsibility to preserve information** and make it available to a **designated community**

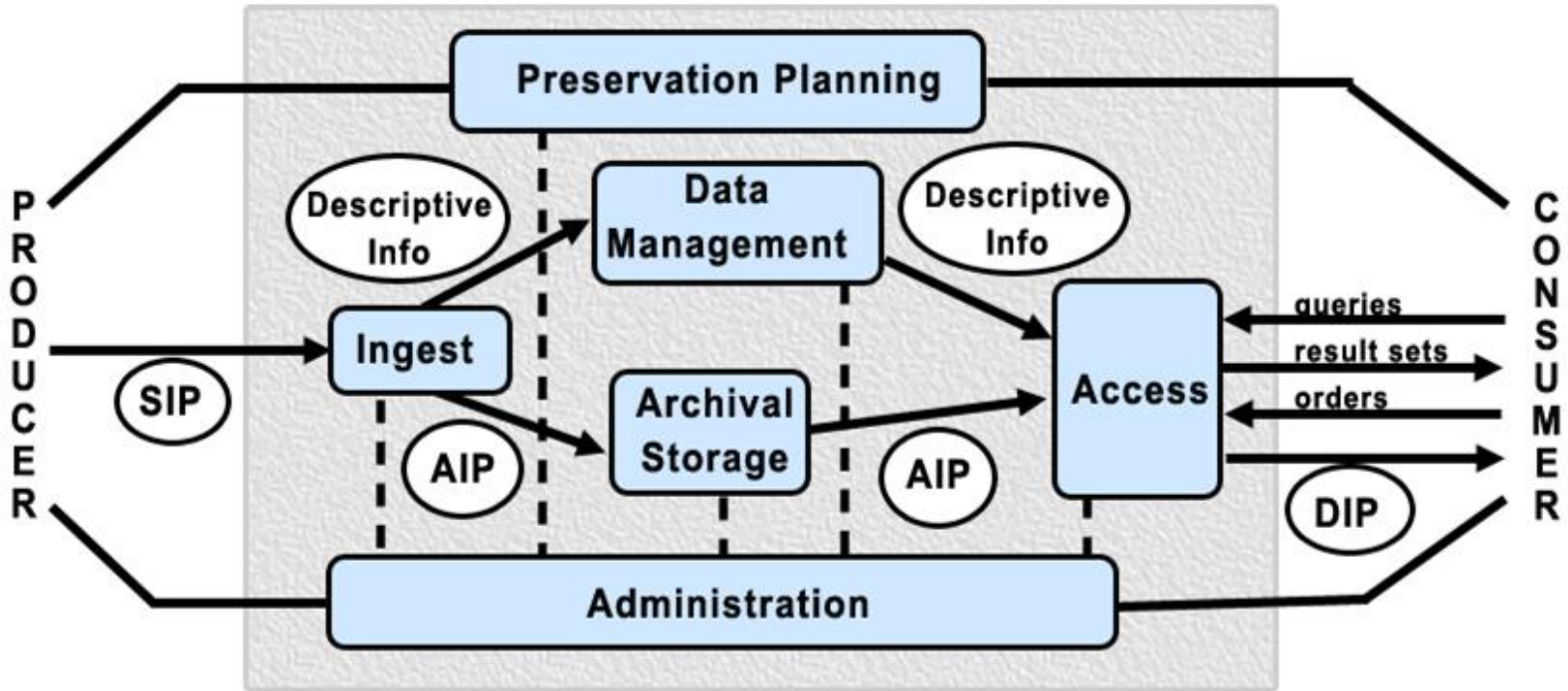


Scary OAISS Spaghetti Monster



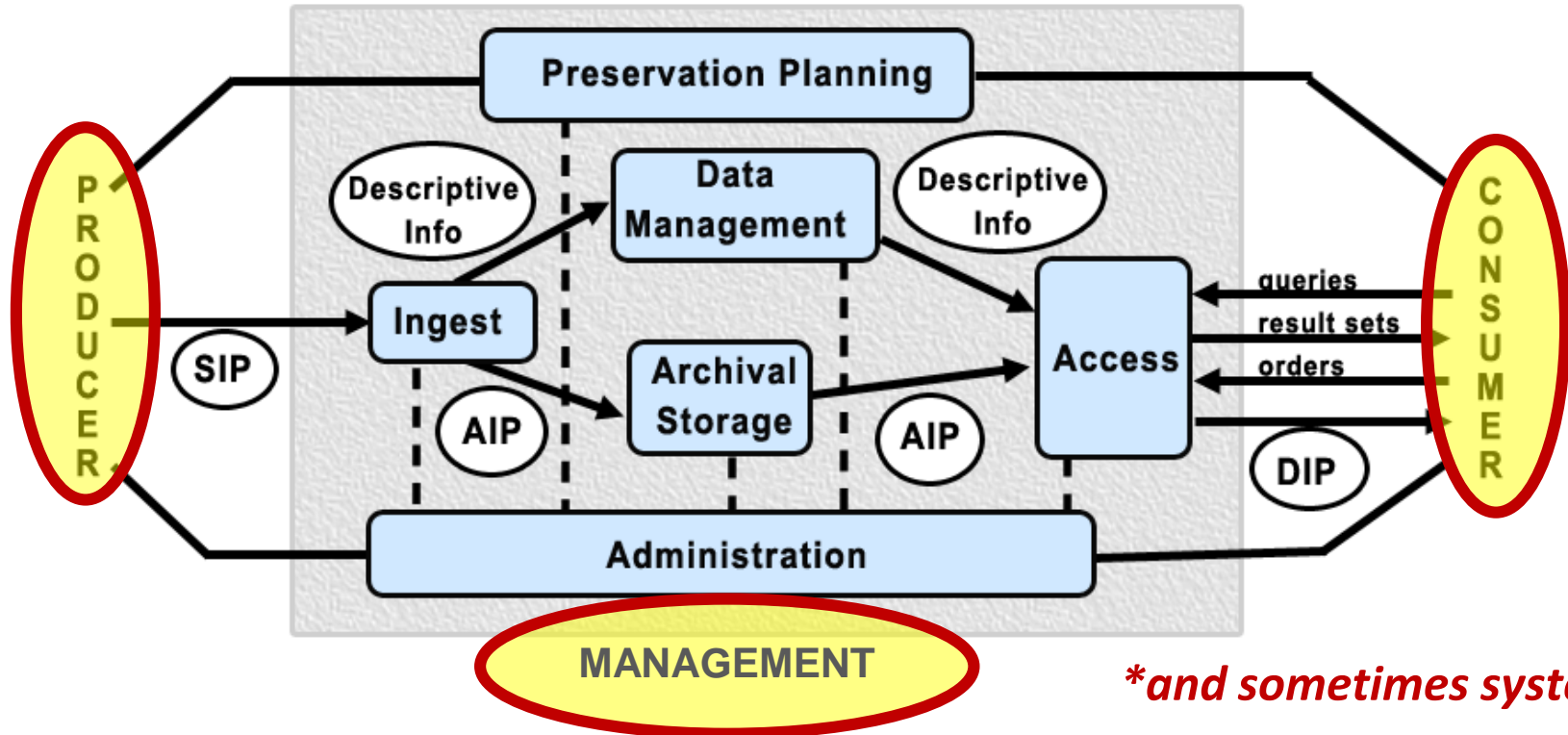
Functional Model...

...still scary but let's give it a chance.



Actors....

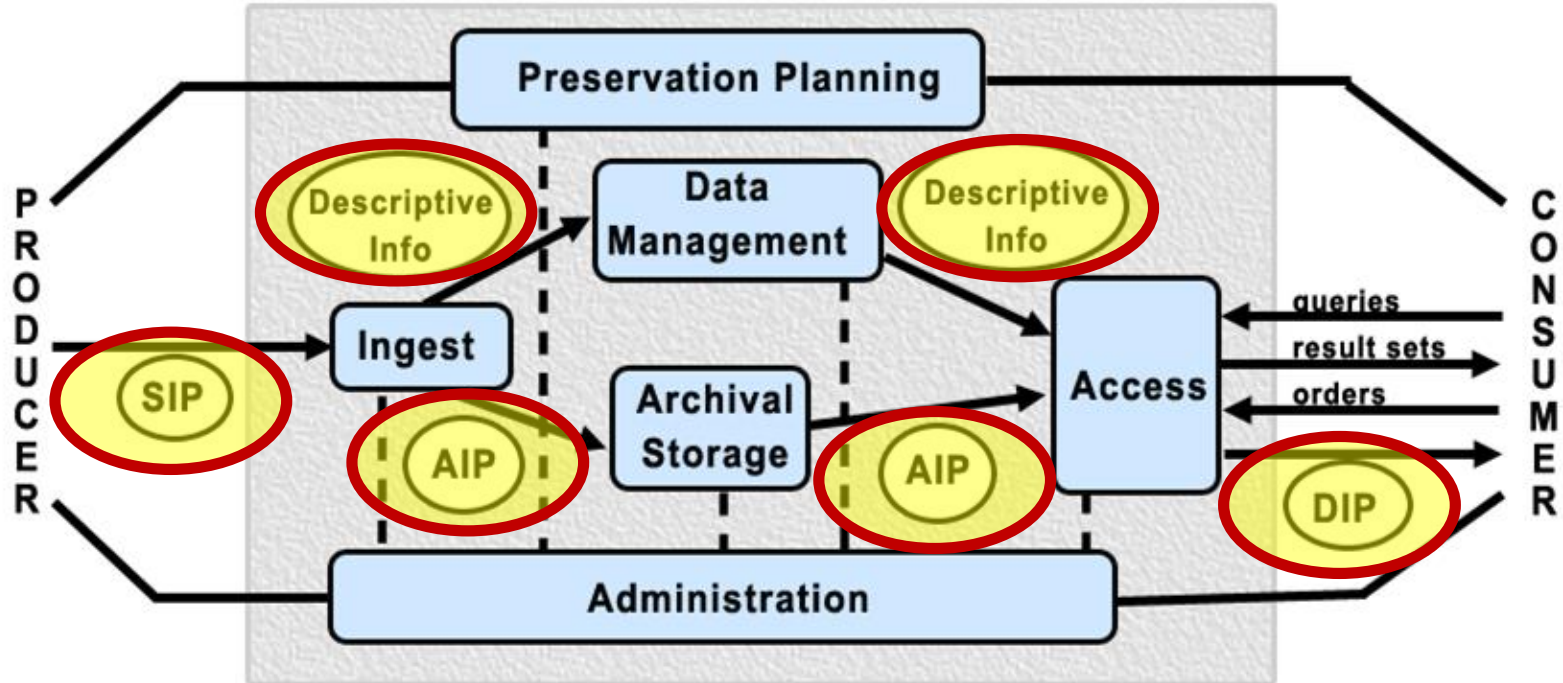
...are just the folks* in your normal professional encounters.



**and sometimes systems*

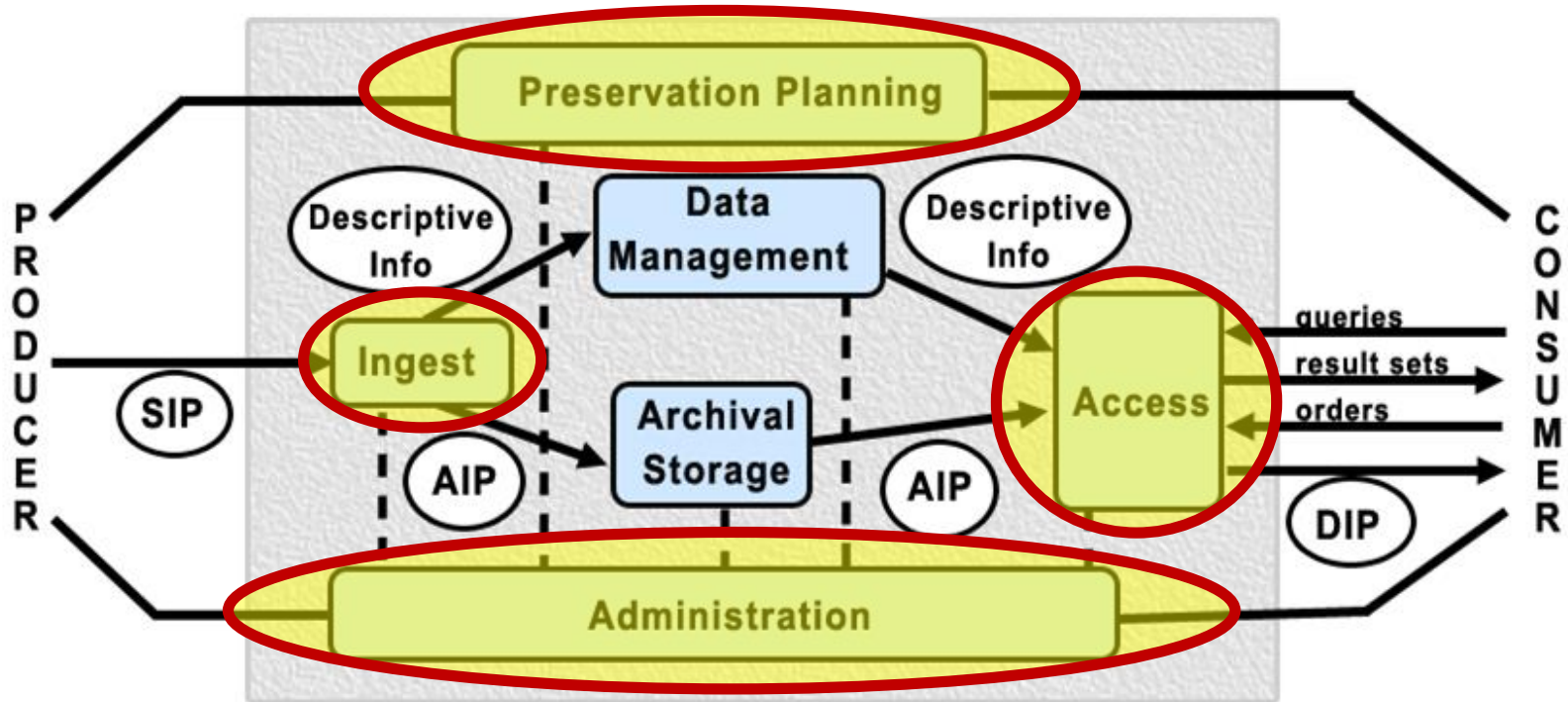
Objects....

...are just the materials and the information about them that bounce around your world.



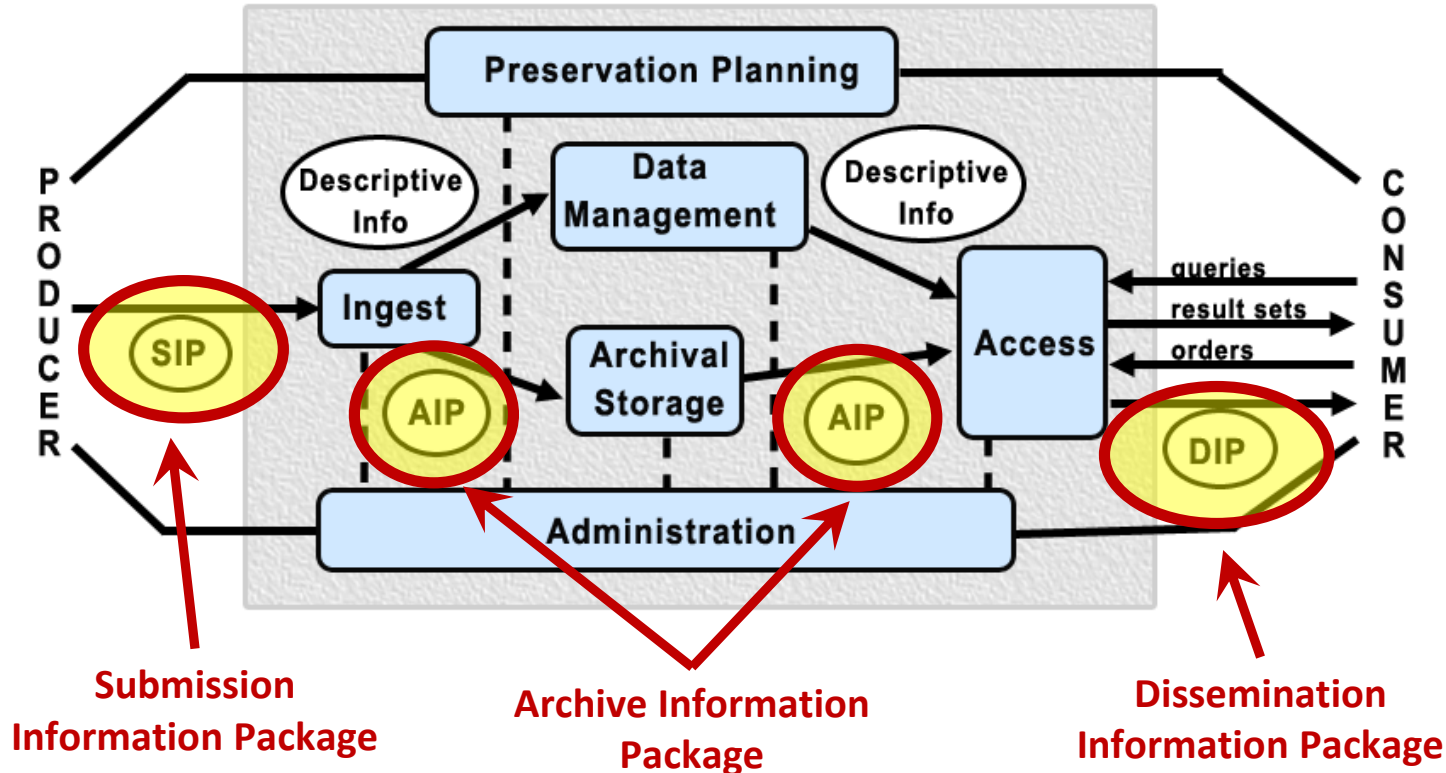
Functional Entities....

...are just the activities that *someone* needs to do in your world.



Information Packages....

...are just a way to keep the materials and the necessary information about them together.



Information Package Structures

May be influenced by:

- Designated community needs
- Existing systems
- Resources available
- Preservation plans

Options from simple to complex

- Standard folder system
- Databases
- XML wrappers

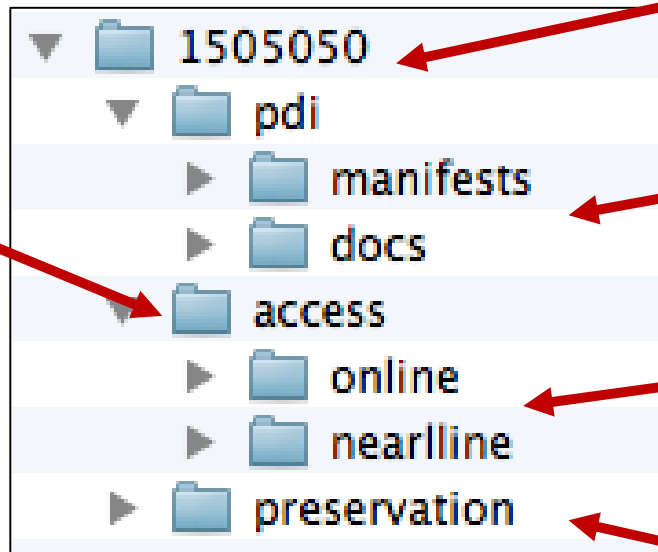


www.digitalbevaring.dk

Tools available to help with creation of IPs

What's in the AIP? An Example

AIP Example from Chris Prom at UIUC Archives



Unique ID
Accession #
System ID

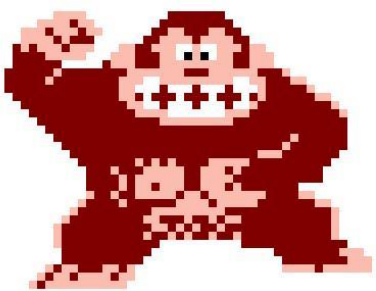
Descriptive info

Online = online version
Nearline = made available in person

Original submitted digital files, pre-preservation actions

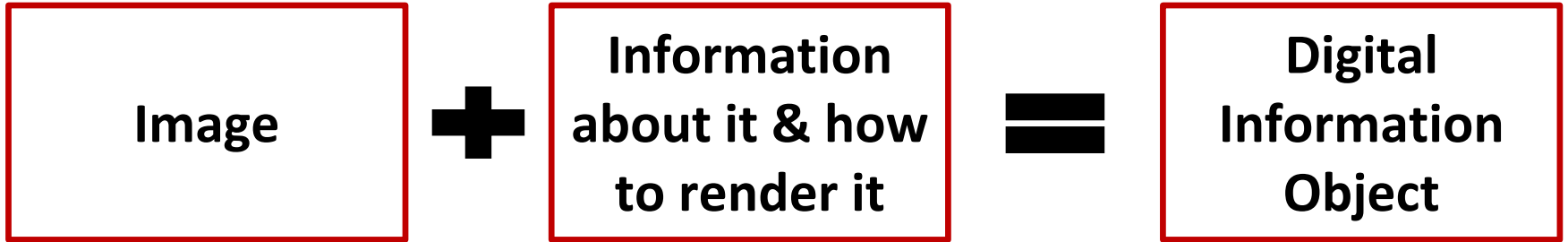
Access copies of original digital files, maybe migrated to new formats

Planet of the AIPs



Getting From Objects To Information

DIGITAL



PHYSICAL



Representation Information

➤ Two types:

Structure Information

File Format, Software....**how to render it!**

Semantic Information

User Documentation, Data Dictionary....**the information about it!**

➤ Can be simple through to very complex

➤ Determined by needs of your Designated Community

➤ Tends to become more complex over time



Preservation Description Information (PDI)



www.digitalbevaring.dk

Supports preservation,
authenticity and dissemination

Describes 'the past and present
states of the data

Consists of 5 components:

- Reference information
- Context information
- Provenance information
- Fixity information
- Access Rights information

A Little Bit On PREMIS

- Widely adopted preservation metadata standard.
- Covers elements of representation information and preservation description information.
- Output is NOT created by hand; depends upon the output of tools who perform actions on your files.
- Record can grow over time, as preservation actions occur.
- Various repository platforms; Archivematica, DataAccessioner and other tools/systems will create PREMIS records for you.

What does PREMIS capture?

PREMIS can capture:

- The program on which the file was created.
- The version of that program.
- The operating system on which that program ran.
- Who created the file.
- The rights associated with the file.
- When the file was ingested into the preservation system.
- Dates the file was validated.
- And more....

METS

Standard for packaging.

Wrapper for XML metadata – you put PREMIS, Dublin Core, MODS, etc **INSIDE** it.

Contains seven sections:

- Header
- Descriptive Metadata
- Administrative Metadata
- File Section
- Structural Map
- Structural Links
- Behavior

METS and PREMIS cover *most* of the metadata requirements of OAIS.

```
<mets>  
  <metsHdr/>  
  <dmdSec/>  
  <amdSec/>  
  <fileSec/>  
  <structMap/>  
  <structLink/>  
  <behaviorSec/>  
</mets>
```



Preservation Models & Packages

QUESTIONS?