

Digital Preservation – Who? What? How? Why? When? Where?

David Giaretta

david@giaretta.org

www.giaretta.org, and www.iso16363.org

2nd Common Preservation Digital Heritage

Lisbon 20/21 June 2016

Lots of questions

- **Who** plays a role in digital preservation?
- **What** is the range of types of digital objects we might need to preserve in future?
- **How** can we make choices about preservation techniques?
- **Why** should society pay to preserve?
- **When** do we need to start preparing for preservation – Data management Plans?
- **Where** should one entrust our digital objects – the Trustworthy Digital Repositories?

Fundamental question: What is digital preservation?

- Paper world - be able to display/ print/ render?
 - But would anyone want to print out the data from CERN?
- Bit world – keep the bits?
 - But a sequence of bits could mean anything!
- OAIS:
 - **Long Term Preservation:** The act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term.

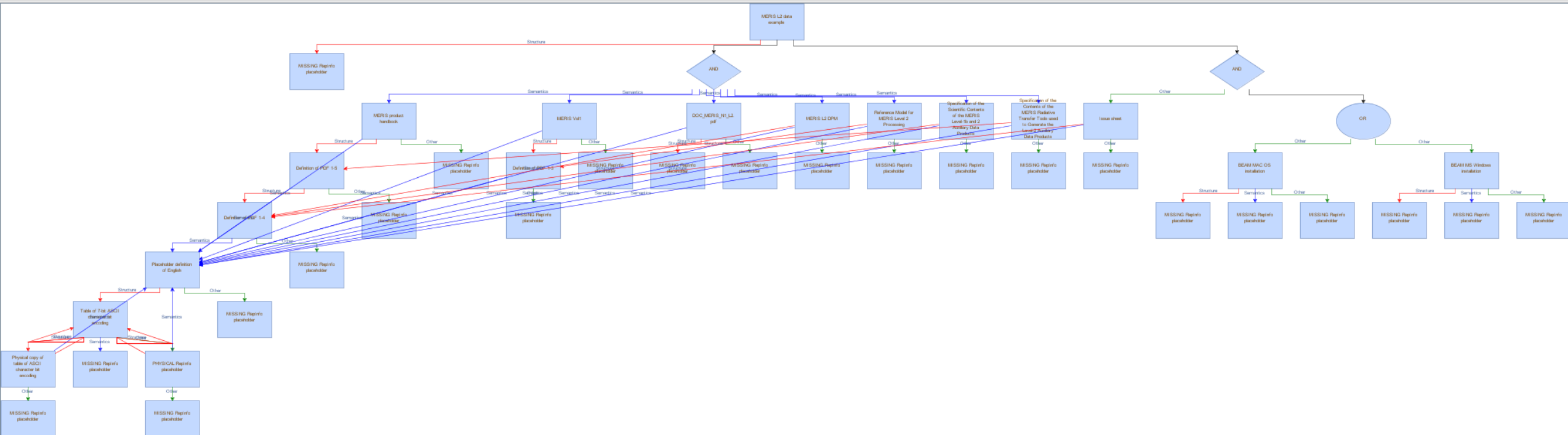
01001110 01001101
01010001 01001101
01010000 01001010
00100000 00100000

Bits...

- We are familiar with just picking up a book and reading it
 - Even very old books
- BUT what if I could just read the bits?
- BITS: 01001110 01001101 01010001 01001101
01010000 01001010 00100000 00100000
- HEX: 4e 4d 51 4d 50 4a 20 20
- Two IEEE 754 32 bit real numbers:
8.6116461E8 1.35644119E10
- Two 32 bit integers
164211241 168379396
- Actually...
....
- ASCII Characters: NMQM PJ
- My flight reference

...semantics ...

Latitude	Longitude	Ozone	Date
132	50	34.9	12/03/1999T17:20:43.1
178	50	45	12/03/1999T19:37:52.7
190	50	78	12/03/1999T21:16:23.9



Possible methods?

- Is just keeping many copies enough?
 - Can we rely on software vendors to solve the problem?
 - Is there a 'magic' preservation form e.g. XML?
 - Is there a special medium that will take care of things forever or at least for hundreds of years?
 - Save the obsolete hardware and software
 - Emulate the hardware
 - Keep migrating
 - Is there someone who will do it for us
 - Is there a risk free option?
 - In any case it takes resources so is there a single figure to tell us how much it will cost?
- **NO**
 - **NO**
 - **NO**
 - **NO**
 - **LIMITED**
 - **RESTRICTIVE**
 - **LOSS OF INFO**
 - **MAYBE**
 - **NO**
 - **NO**

Techniques

- Emulate or migrate is often heard
 - Too limited – what about SEMANTICS etc?

- BETTER:

- Add Representation Information
- Transform
- Hand over



OAIS introduced “Transformational Information Properties” – more general than “Significant Properties”

The OAIS “Archival Information Package” – make sure all the relevant information is ready

Who plays a role in digital preservation?

- Data creators
- Funders of data creation
- Archivists
- Technical support
- Users of the archives
 - Designated Community
- Funders of the archives

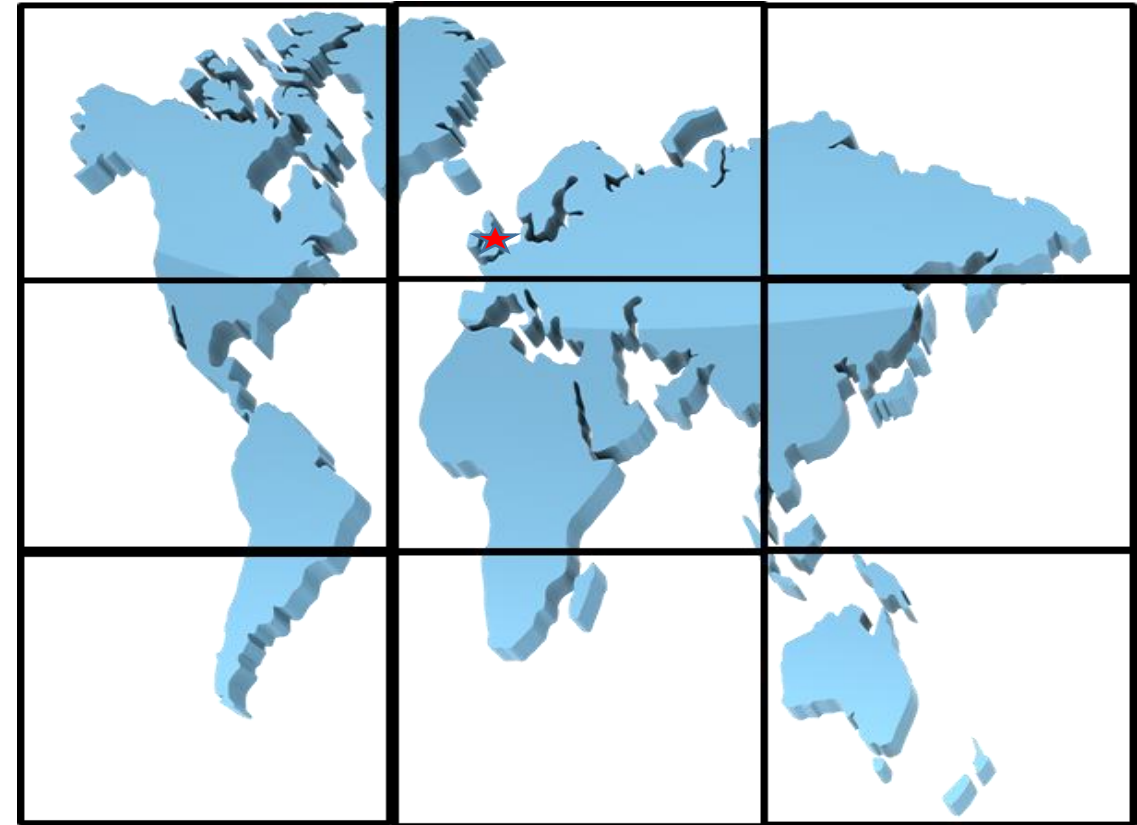
What is the range of types of digital objects we might need to preserve in future?

- Just rendered objects?
 - Rely on a human to interpret the rendering
 - One human ⇔ reading one book at a time?
- Scientific Data?
 - Needs to be able to
 - be machine processable
 - combine one dataset with another
 - make sure we do not add “apples” with “oranges” or “miles” with “kilometres” by accident
 - One machine ⇔ combining hundreds of datasets at a time
 - NEED EXPLICIT SEMANTICS....
 - OAIS: STRUCTURE, SEMANTICS and OTHER

- Simple documents and images
 - Simply display for human viewing
- Scientific data - containing numbers, text etc
 - Not usually for simply viewing by human
 - The meaning of each number is not usually guessable
- Various types of complex objects e.g. databases
 - Many interrelationships between the various elements which affect the meaning
- Changing over time - in the sense that the state at a previous time is important
 - annotations

How can we make choices about preservation techniques?

- Lots of lists of tools – many the same
- Which tool/technique to try first?
- Need a map / co-ordinates to tell where we are and which way to go



For digital objects – one suggestion

- Rendered vs non-rendered
 - scientific data vs simple documents
 - Static vs dynamic
 - objects that don't change vs objects that do (such as growing lists)
 - Simple vs composite
 - simple document vs documents with linked spreadsheets
 - simple documents vs data made up of multiple different components
 - Passive vs active
 - word file vs word application
-
- **These can be combined to give a set of combinations e.g. non-rendered/static/composite/passive**
 - **From these one can build a way to help decide on which tool/technique to use for a given digital object**

Why should society pay to preserve?

- For future generations?
 - ...but
- Everything?
- What does it cost?
- Who pays and Why?
 - **VALUE**
 - Now ... or ..
 - In the future (potentially)
 - How to increase value?
 - **Increase usability – so it is good that OAIS links preservation to usability!**

Who knows?

- The people who create the digital objects
- For scientific data
 - The creator's understanding of the data may change over the project e.g.
 - How is the data calibrated?
 - What instrumental effects need to be taken into account?
 - How is this part of the data related to that part?
 - Other research is published which changes the understanding
 -
 - How was it processed?
 - Who could have changed it?
 - Is it what we think it is?
- But it is a burden to collect all this!

When should Additional Information be collected?

Data Management Plans (DMP)

- Funders are asking for Data Management Plans
 - BUT..
 - Short
 - Few details
 - Written at the start and usually not revised
 - Usually plan implementation is NOT checked
 - SO...
- ACTIVE DATA MANAGEMENT PLANS (ADMP)
 - Group who wrote OAIS and ISO 16363 are preparing a standard: ***Information Preparation to Enable Long Term Use***
 - Specify “Additional Information” needed at various times through the “project”
 - RDA ADMP group is holding an ***International and Interdisciplinary workshop on Active Data Management Plans*** <https://indico.cern.ch/event/520120/> at CERN 28-30 June
 - Work on processes and tools to capture Additional Information

CCSDS Standards

Guides the collection/creation of the Additional Information required

**Information
Preparation
to Enable
Long Term
Use**

Defines a mechanism to transfer Data and the appropriate Additional Information to the archive

PAIMAS

PAIS

Defines how the information should be preserved

OAIS

Defines how to check that the information is being preserved

**AUDIT &
CERTIFICATION**

EAST

DEDSL

PVL

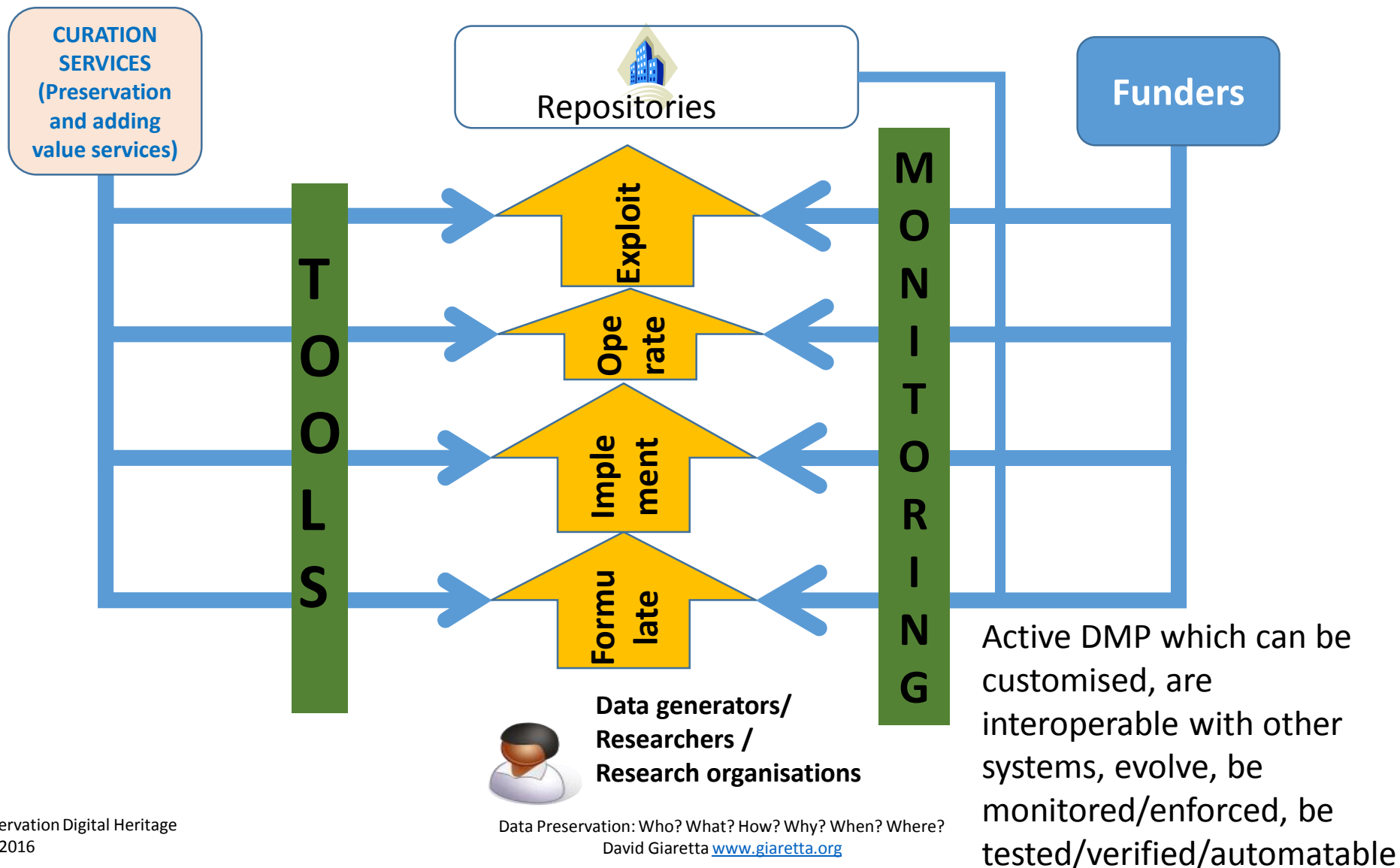
XFDU

CCSDS/ISO Standards for Representation
Information

CCSDS/ISO Standard
for Packaging

Many other standards could be used for Representation
Information

Research Data Alliance – Active Data Management Plan group



Where should one entrust our digital objects – the Trustworthy Digital Repositories?

- An organisation (A) says it are doing a good job in preservation:
 - Do you believe it?
 - Would you entrust your valuable information to it?
- Suppose another organisation (B) told you that organisation (A) was doing a good job (and (A) said (B) was also doing a good job)..
 - What is being checked?
 - Same questions
- You would probably demand **evidence!**

Do not believe all claims!

ABOUT ∨ SERVICES ∨ FACTS & FIGURES ∨ FEE

Why a doctorate degree?

Although quite demanding, people would still prefer to reach for greater reason why a doctorate is being offered to maximize the potential medical doctor by degree is such music in the ears. To buy a doctorate possible academic degree, you can explore a lot of opportunities if you have

GET YOUR DEGREE

Buy a PhD to Get the Experience

If you **buy a PhD** you will achieve promotions at your workplace without will ruin your family or work life. If you **buy a PhD** from our company with the respect of your employers and co workers.

If you **buy a PhD** from our company you will purchase an accredited reputation we collaborate with accredited universities we help you **buy a PhD** with great paid jobs. All you have to do is to fill the online application in order to detail to provide the work tasks that you have completed. You can provide additional the institution that you are eligible to **buy a PhD**.

We check the documents that you have submitted in order to verify if they are relevant and genuine. You should stay connected with our adviser in order to update any error or additional requirements with our application. You will receive transcripts that can be used to prove that you have been qualified from a prestigious university. When you

How To Make Someone a Saint!

(or call 818-233-0262 for assistance!)

Order Now

You've probably never made someone a saint before. Don't worry, not many people have. We've made the process very easy. It takes a little more effort than buying a book on the Internet because this is a very personalized gift.

You can either go directly to the [order form](#), or use the guide below. This guide will step you through the ABC's of sainthood. As you'll see, we provide lots of help along the way. If you want, you can print this page and use it as a guide as you go along.

Here's the overview:

- A** First - You [get ready](#) by deciding some basics.
- B** Then you [work out the details](#) of sainthood for your saint.
- C** Finally, you [place the order](#).

If you would prefer to fill out an order form to fax or mail us, click here for an [alternate order form](#).

A Getting ready involves three simple steps:

1 Pick a person and occasion.

The person could be your husband, wife, friend, relative, co-worker or anybody you're close to. The occasion could be a birthday, holiday, special event like a wedding or graduation, or "just because".

2 Decide what you want to give them.

Basically there are two options: a Sainthood Certificate and a Cyber Shrine. The certificate is a beautiful document that makes a wonderful present. But if you don't need or want the document itself to give to your saint, you can pick a Cyber Shrine. The best is to do both so your saint can not only admire the certificate but can visit their own shrine on the Internet and let others visit it whenever they want. [View the options](#) and then return here by pressing the back button on your browser.

3 Photo or not?

Having a photo of your saint with a halo makes the gift much more special and personal, but if you don't have a photo, we have a St. Cyber (a generic saint) we'll use instead. Learn about the [photo options](#). If you decide to use a photo, round one up so you'll be ready to get it to us.

B Once you've made your decisions, you're ready to start the saint-making process. To help with the personalization information, we've provided the form that follows. It's just like the Sainthood Nomination form you'll use to order. We've even put in some [Easy Steps](#) in green. These are suggestions you can follow in case you get stuck. But don't feel you're limited to the [Easy Steps](#) suggestions. Use your imagination and have fun with this!

1 Pick a Saint Name

This is usually something like "Saint Mildred of Canton". Using a "place" will distinguish one saint from another. The "place" can be a city, street or neighborhood - or a company or organization such as "Saint Carl of the Mounted Police".

Write the choice you've selected below:

Saint _____ of _____

Where should one entrust our digital objects – the Trustworthy Digital Repositories?

- An organisation (A) says it are doing a good job in preservation:
 - Do you believe it?
 - Would you entrust your valuable information to it?
- Suppose another organisation (B) told you that organisation (A) was doing a good job (and (A) said (B) was also doing a good job)..
 - What is being checked?
 - Same questions
- You would probably demand **evidence!**
- ISO 16363 (see www.iso16363.org)
 - we depend on ISO for so much in our lives – food, health, cars, planes etc

Useful Links

- OAIS
 - WIKI pages: wiki.oais.info
 - WEB pages: www.oais.info
 - Ideas for OAIS updates in 2017:
<http://wiki.oais.info/bin/view/Bugs/WebChanges>
- ISO 16363:
 - www.iso16363.org
- Active Data Management Plans:
 - CCSDS/ISO
 - http://cwe.ccsds.org/moims/default.aspx#_MOIMS-DAI
 - Research Data Alliance:
 - <https://www.rd-alliance.org/groups/active-data-management-plans.html>
- Me:
 - www.giaretta.org



END

david@giaretta.org