What is metadata?

Metadata structure

Types of metadata

Conclusion

Metadata

FOSTER

Discovering Open Practices for Early Career Researchers
4 September 2014

Laurence Horton
Data Librarian, The London School of Economics and Political Science
L.Horton@lse.ac.uk

This presentation is licensed under a Creative Commons Attribution 4.0 International license.

Prezi

What is metadata?

Metadata structure

Types of metadata

Conclusion

Metadata

FOSTER

Discovering Open Practices for Early Career Researchers
4 September 2014

Laurence Horton
Data Librarian, The London School of Economics and Political Science
L.Horton@lse.ac.uk

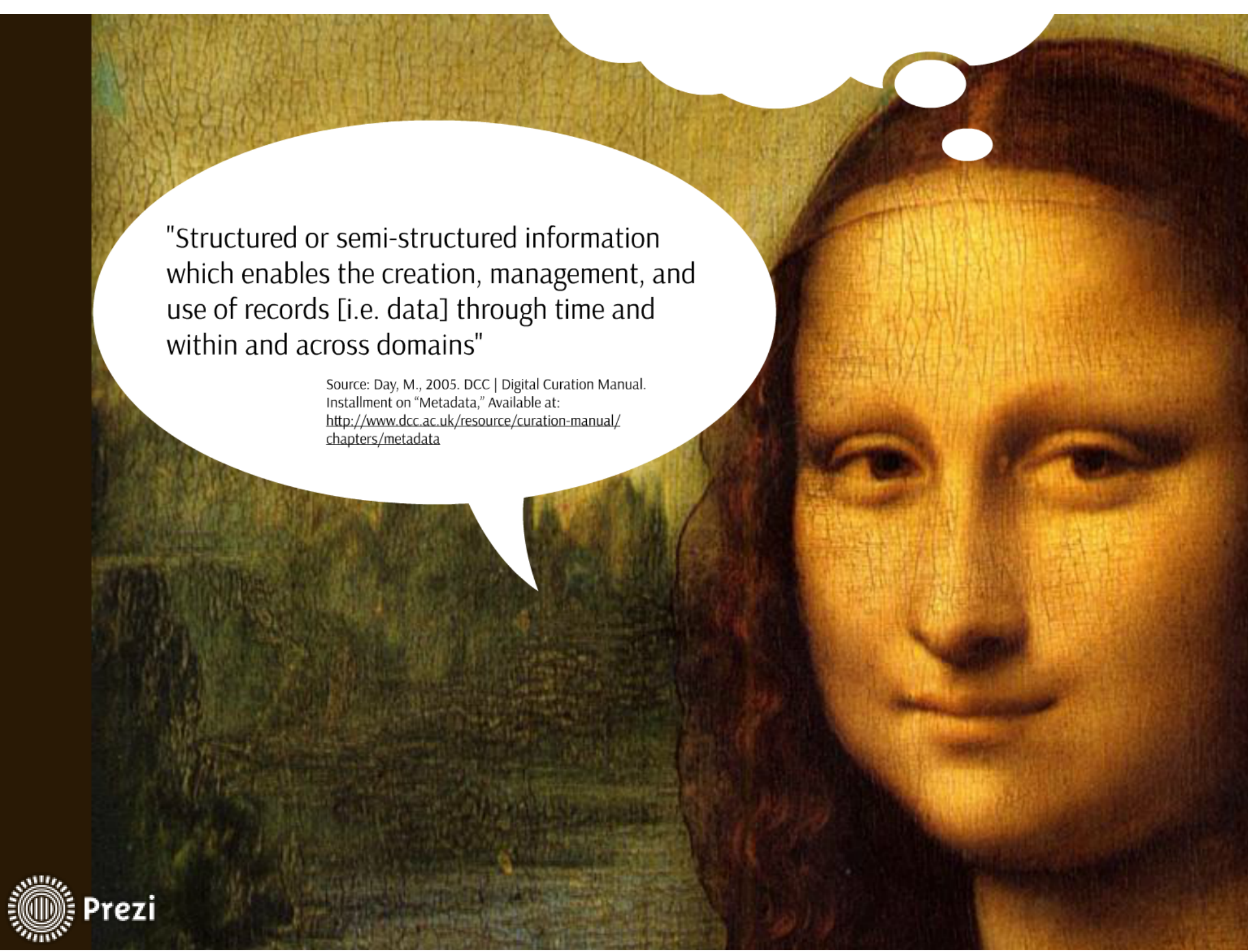
This presentation is licensed under a Creative Commons Attribution 4.0 International license.

Prezi

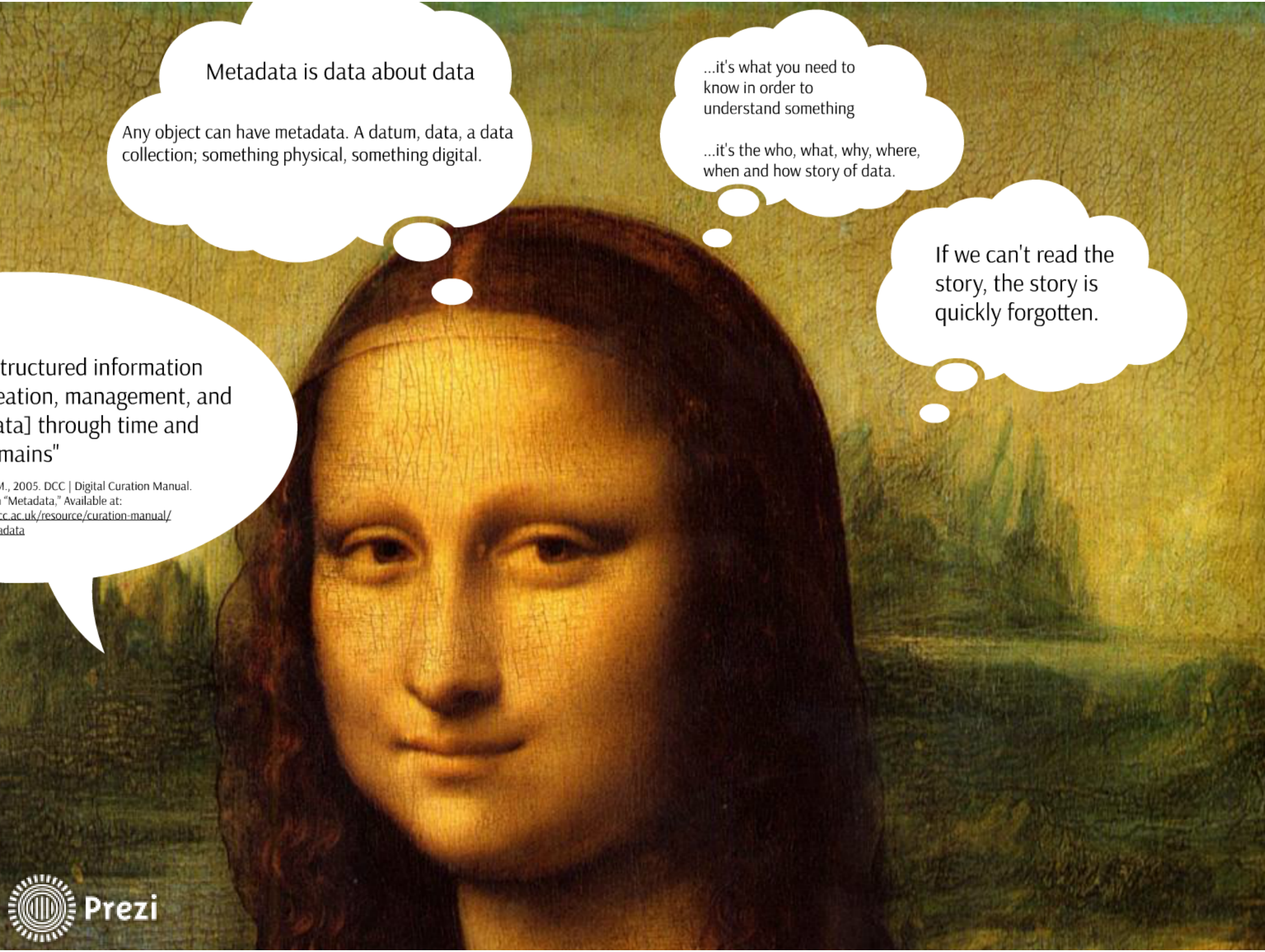
Metad

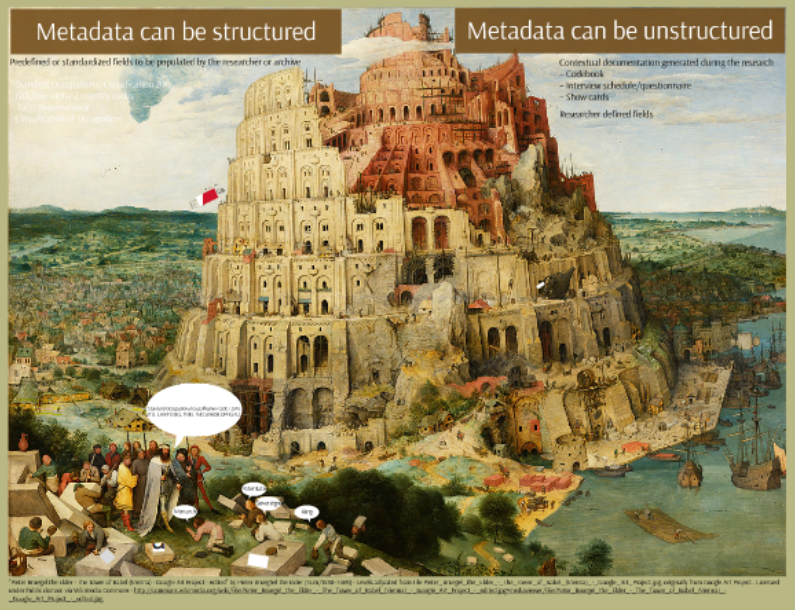
What is metadata?

Prezi


"Structured or semi-structured information which enables the creation, management, and use of records [i.e. data] through time and within and across domains"
Source: Day, M., 2005. DCC | Digital Curation Manual. Installment on "Metadata," Available at: http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata

Prezi


Metadata is data about data
Any object can have metadata. A datum, data, a data collection; something physical, something digital.
...it's what you need to know in order to understand something
...it's the who, what, why, where, when and how story of data.
If we can't read the story, the story is quickly forgotten.
tructured information
eation, management, and
ata] through time and
mains"
M., 2005. DCC | Digital Curation Manual.
"Metadata," Available at:
cc.ac.uk/resource/curation-manual/
adata
Prezi



Metadata structure

Prezi

Metadata can be structured

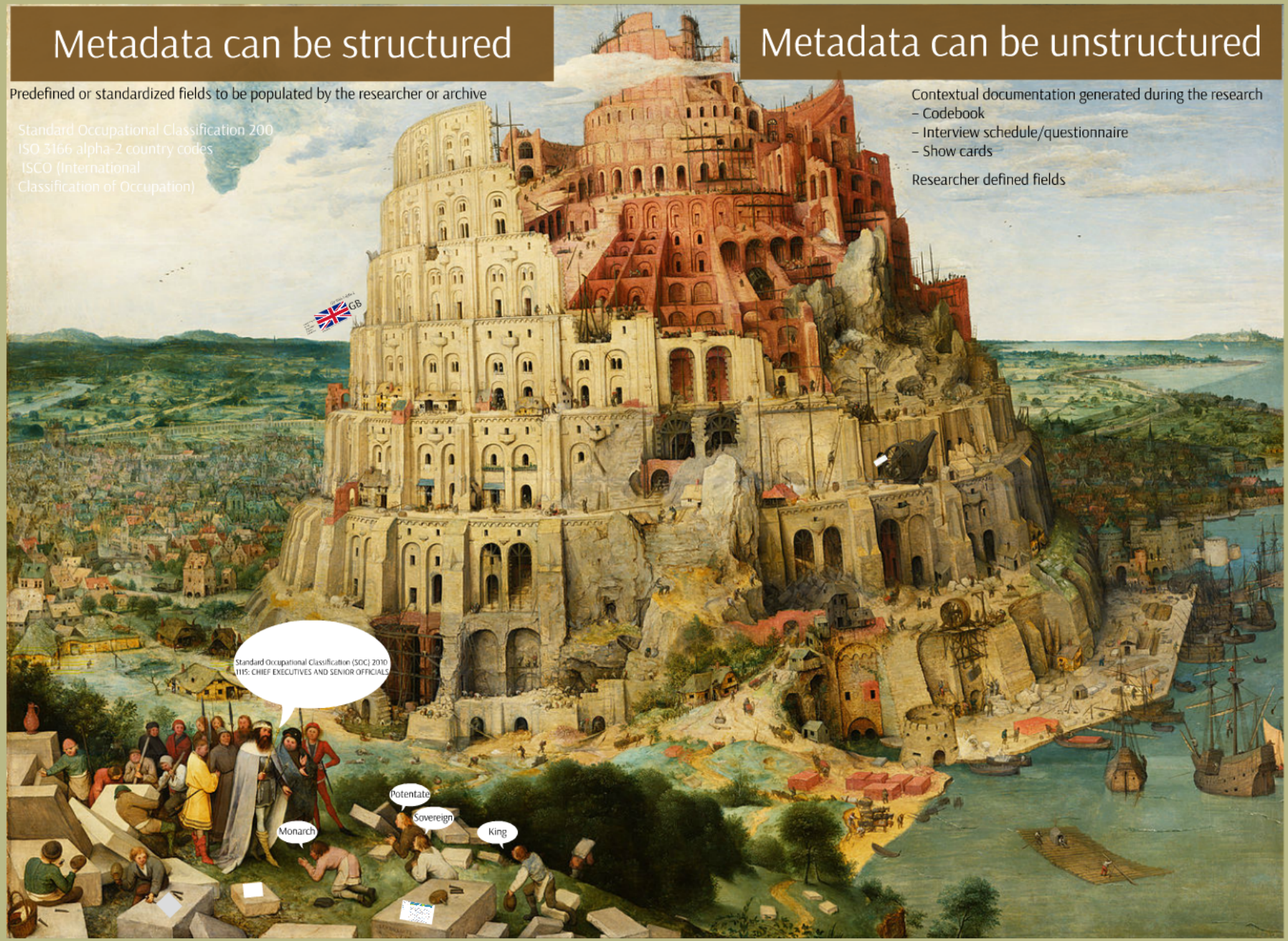
Predefined or standardized fields to be populated by the researcher or archive

Standard Occupational Classification 200
ISO 3166 alpha-2 country codes
ISCO (International
Classification of Occupation)

Metadata can be unstructured

Contextual documentation generated during the research
- Codebook
- Interview schedule/questionnaire
- Show cards

Researcher defined fields



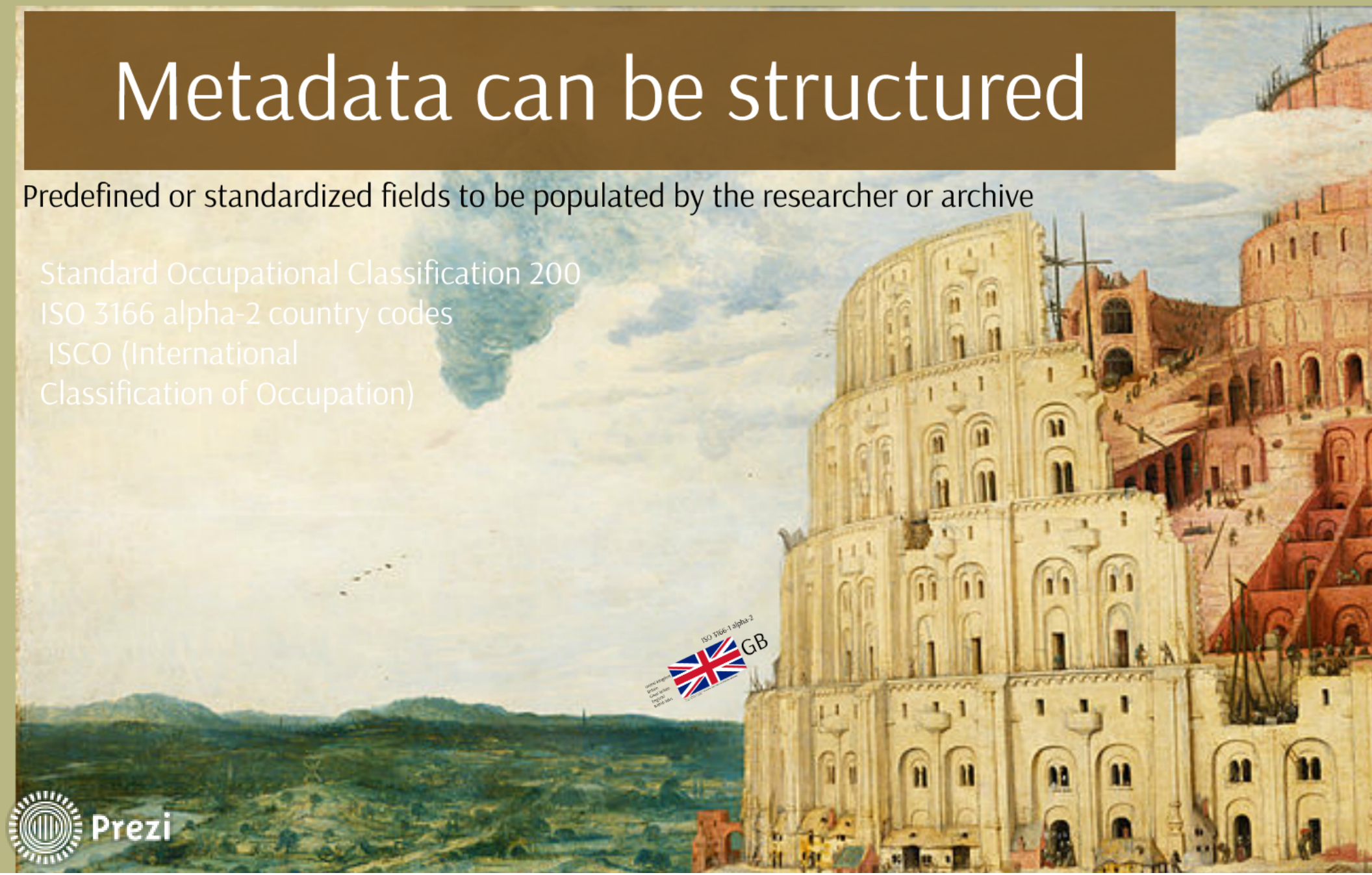
"Pieter Bruegel the Elder - The Tower of Babel (Vienna) - Google Art Project - edited" by Pieter Brueghel the Elder (1526/1530–1569) - Levels adjusted from File:Pieter_Bruegel_the_Elder_-_The_Tower_of_Babel_(Vienna)_-_Google_Art_Project.jpg, originally from Google Art Project.. Licensed [illegible] Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Pieter_Bruegel_the_Elder_-_The_Tower_of_Babel_(Vienna)_-_Google_Art_Project_-_edited.jpg#mediaviewer/File:Pieter_Bruegel_the_Elder_-_The_Tower_of_Babel_(Vienna)_-_[illegible]_Art_Project_-_edited.jpg

Prezi

Metadata can be structured

Predefined or standardized fields to be populated by the researcher or archive

Standard Occupational Classification 200
ISO 3166 alpha-2 country codes
ISCO (International Classification of Occupation)



Prezi

ISO 3166-1 alpha-2

United Kingdom
Britain
Great Britain
England
British Isles

"Flag of the United Kingdom". Via Wikipedia - http://en.wikipedia.org/wiki/File:Flag_of_the_United_Kingdom.svg#mediaviewer/File:Flag_of_the_United_Kingdom.svg

GB

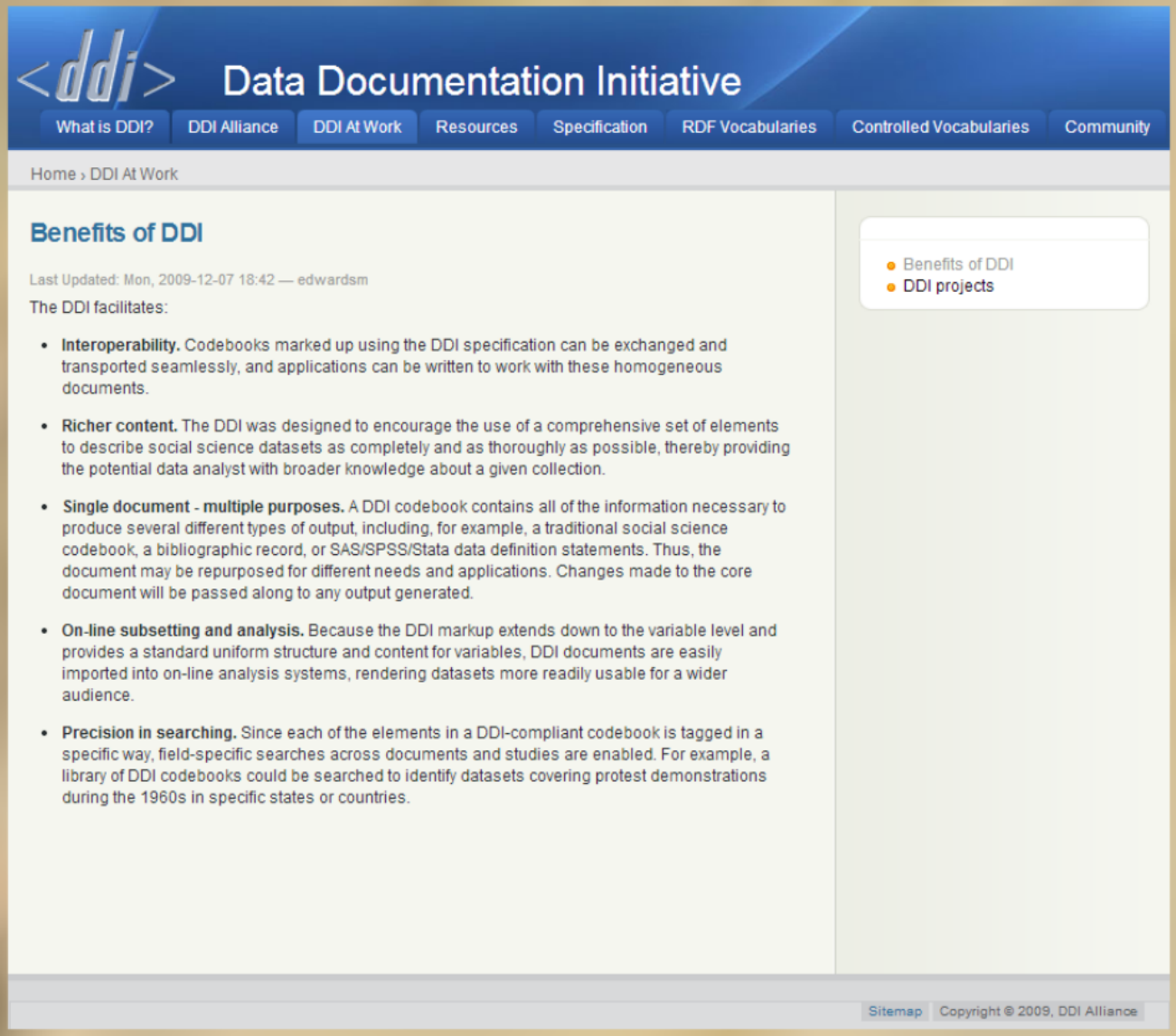
Prezi

Pieter Bruegel the Elder - The Tower of Babel (Vienna) - Google Art Project - edited" by Pieter Brueghel the Elder (1526/1530–1569) - Levels adju
nder Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Pieter_Bruegel_the_Elder_-_The_Tower_of_Babel_
Google_Art_Project_-_edited.jpg

Prezi

der - The Tower of Babel (Vienna) - Google Art Project - edited" by Pieter Brueghel th
Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Pieter_Bruegel_

Prezi



Source: DDI Alliance (2009) "Benefits of DDI" http://www.ddialliance.org/ddi-at-work/benefits

Prezi

gel the Elder - The Tower of Babel (Vienna) - Google Art Project - edited" by Pieter Brueghel the Elder (1
c domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Pieter_Bruegel_the_Elde
rt_Project_-_edited.jpg

Prezi

What is a DOI?

A DOI is a Digital Object Identifier. That is a ***Digital Identifier*** of an ***Object***, not an Identifier of a Digital Object.

10.1234/456-mydoc-456584893489 is an example of the DOI format. It consists of a Prefix and a Suffix. The prefix always begins '10' indicating this is a DOI name; the second part of the prefix is typically four digits, and is allocated to the DOI assigner. The suffix is created by the assigner, and can be any length and incorporate other numbering schemes (including an existing identifier) if desired. Existing data can also be linked in the DOI record.

A DOI can be assigned to any Object. An object can be anything you want to manage: a piece of online content in a particular version, or a related object, such as an author or a party to an agreement; or a physical asset, a DVD or an item of equipment, for example.

DOIs can be assigned at different levels of granularity to suit differing needs, e.g. a collection, a book, a table within the book, the related source data for the table etc. When assigning DOIs, the registrant links metadata with an appropriate level of detail to identify a separate entity within the DOI system.

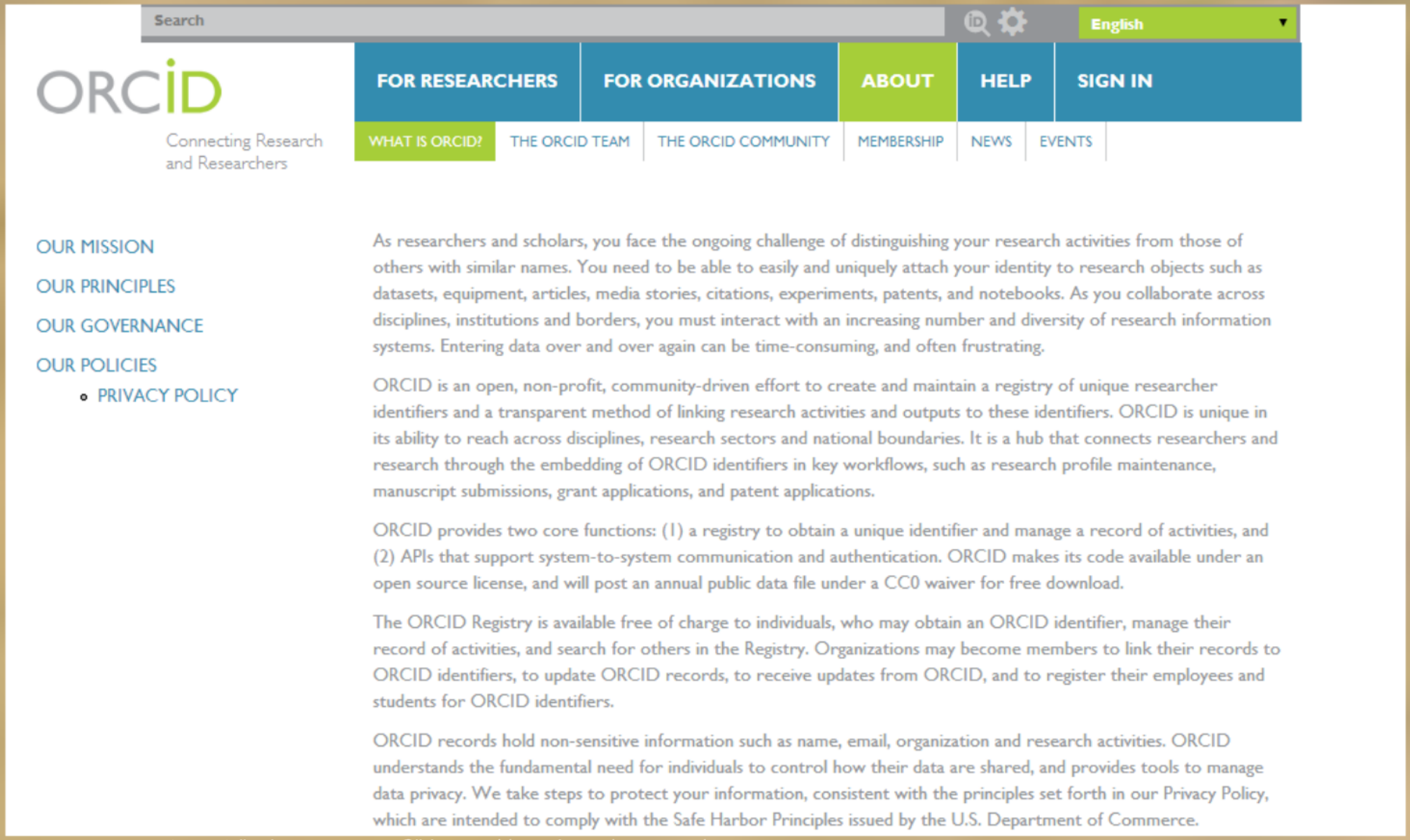
DOIs in use

- First applications launched in 2000
- Currently being assigned by well over 5000 naming authorities e.g. publishers, science data centres, movie studios etc
- Around 100 million DOI names assigned
- Over 1.5 billion DOI resolutions per year

International DOI Foundation (2014) "Driven by DOI®" http://www.doi.org/driven_by_doi/DOI_Marketing_Brochure.pdf

Prezi

Elder - The Tower of Babel (Vienna) - Google Art Project - edited" by Pieter Brueghel the
in the Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Pieter_Bruegel_th
ect_-_edited.jpg



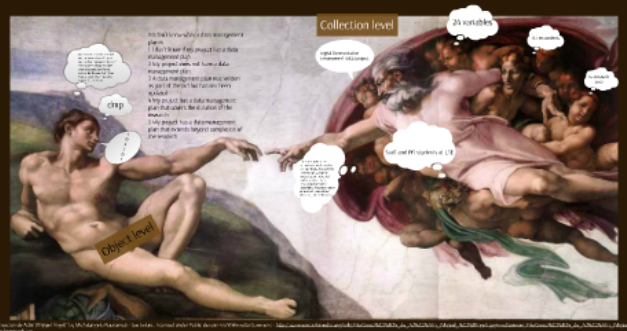
Source: ORCID (2014) "What is ORCID?" https://orcid.org/content/initiative

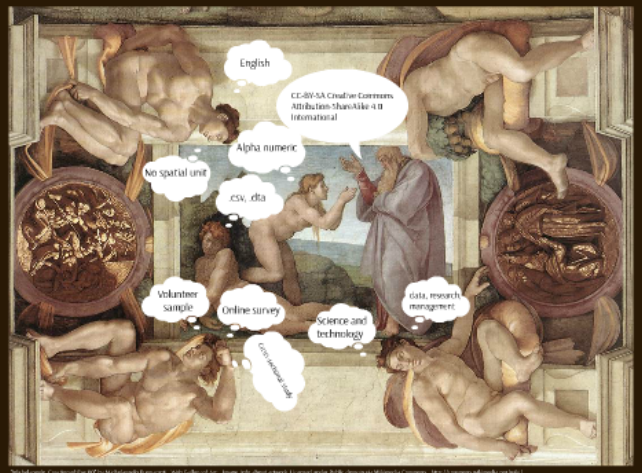
Prezi

Unstructured metadata at work....

Source: Jessica Trixa (2014) "Metadata and documentation" CESSDA Training, licensed under a Creative Commons Attribution 3.0 Unported License.

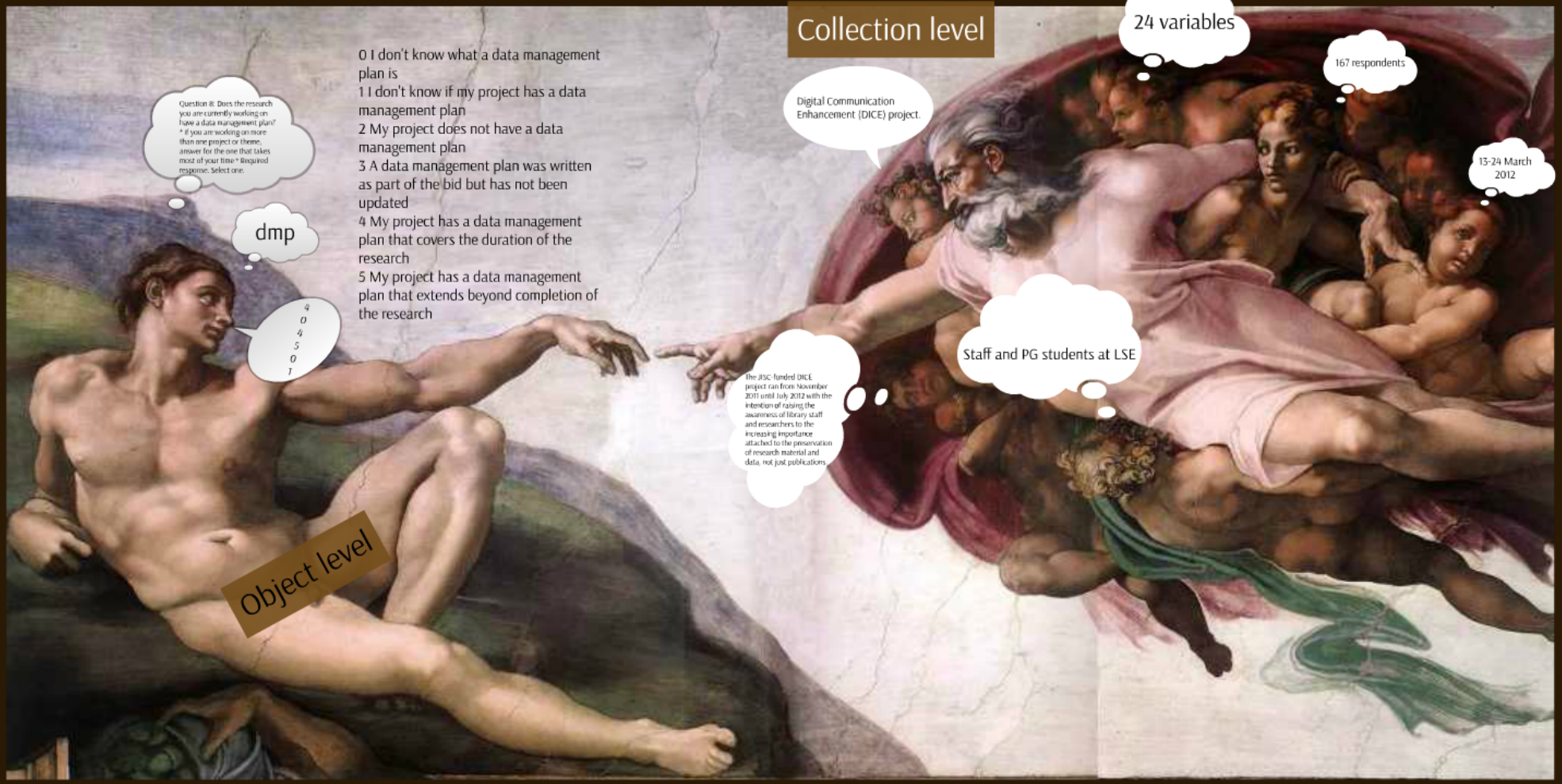
Prezi





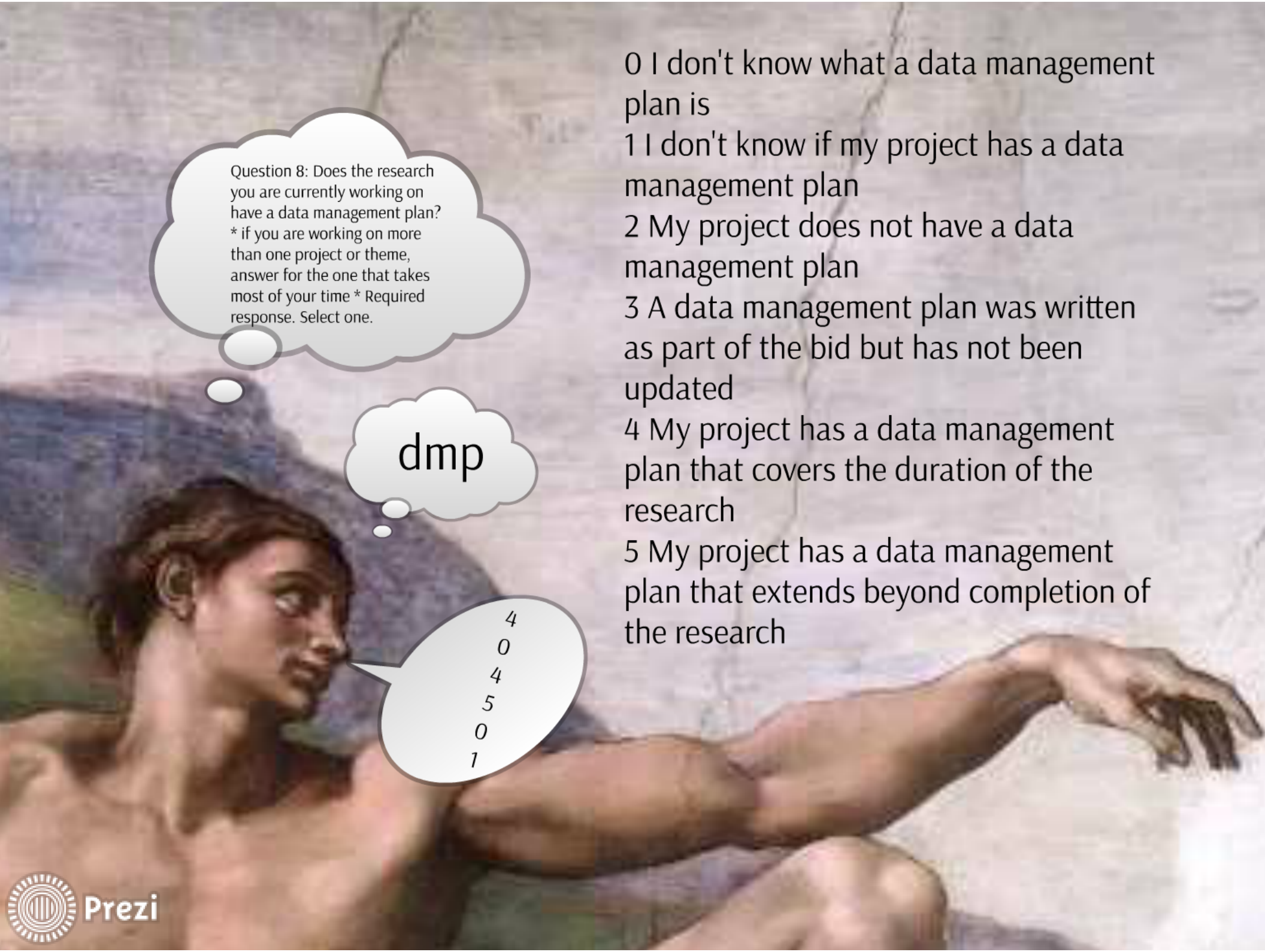
Types of metadata

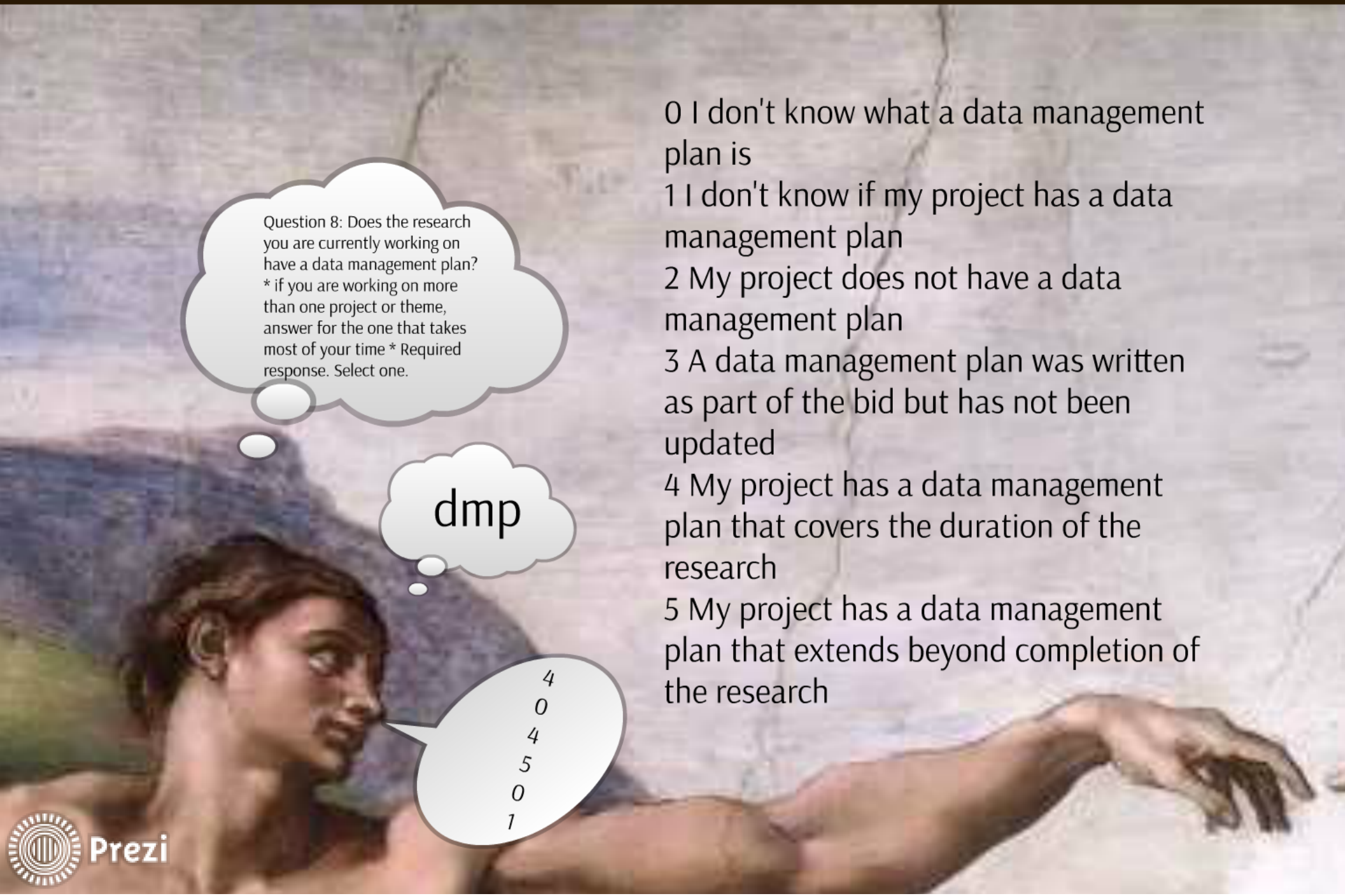
Prezi

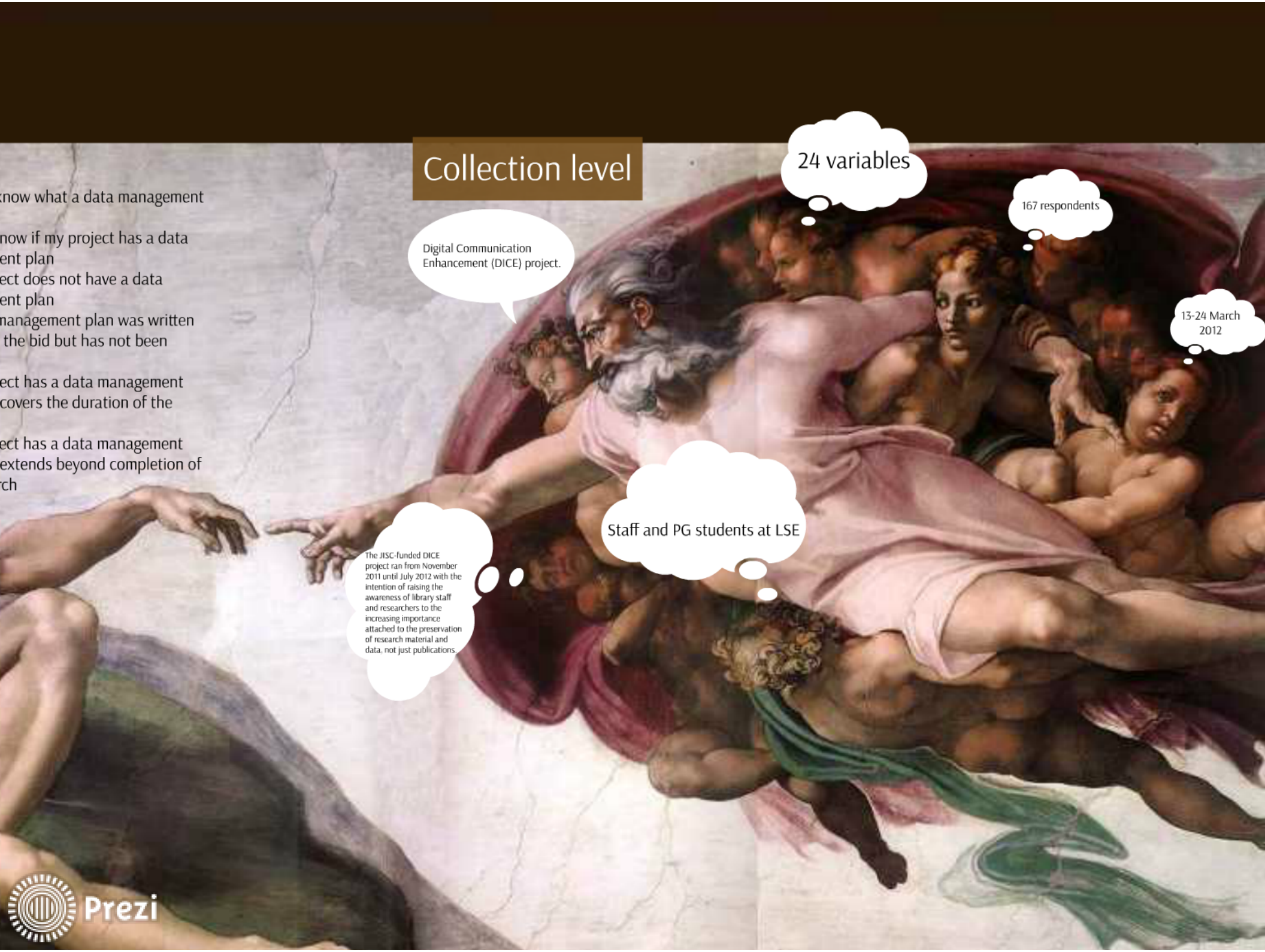


"Creación de Adán (Miguel Ángel)" by Michelangelo Buonarroti - See below.. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Creaci%C3%B3n_de_Ad%C3%A1n_(Miguel_%C3%81ngel).jpg#mediaviewer/File:Creaci%C3%B3n_de_Ad%C3%A1n_(Miguel_%C3%81ngel).jpg

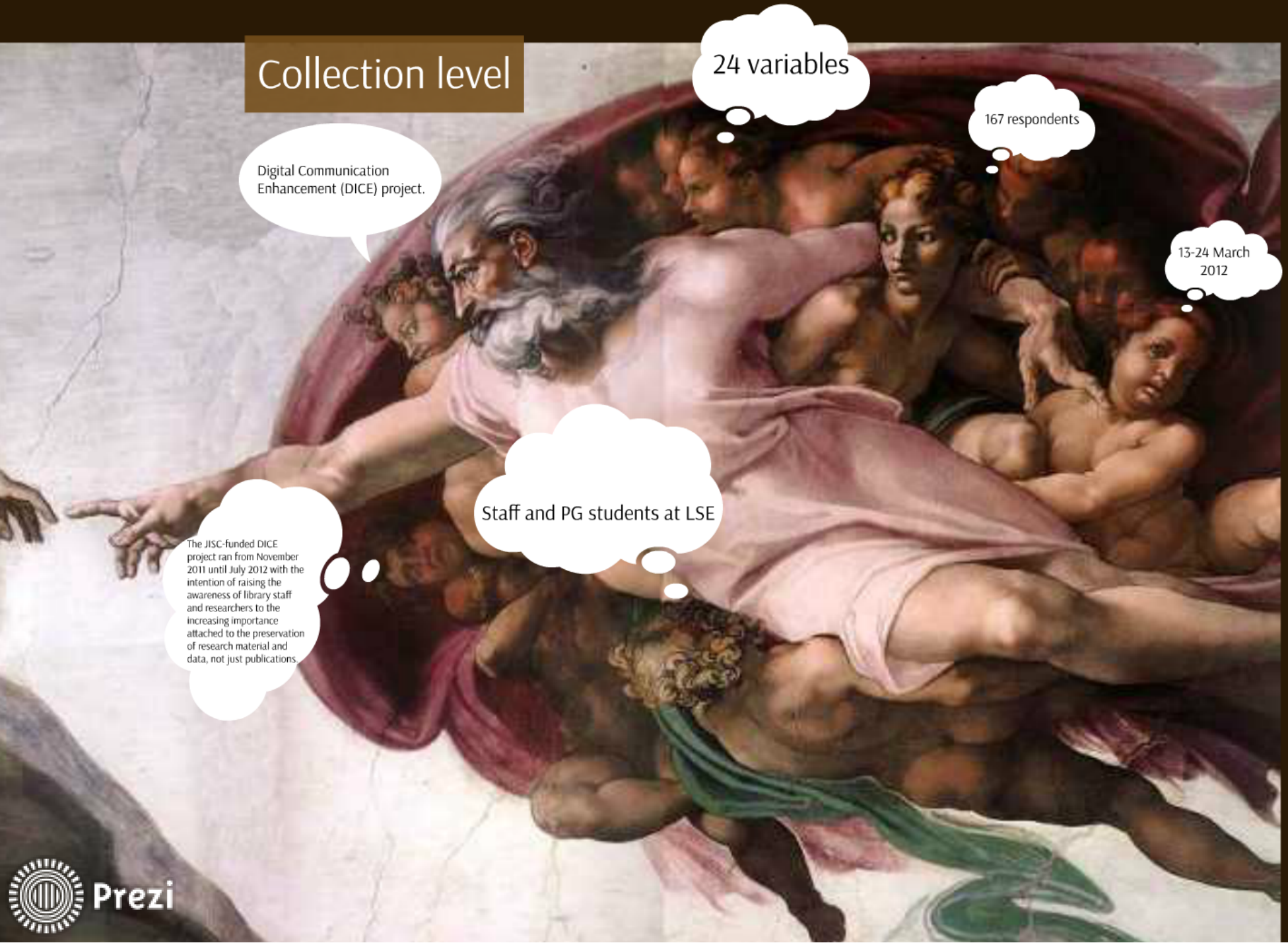
Prezi


Question 8: Does the research you are currently working on have a data management plan? * if you are working on more than one project or theme, answer for the one that takes most of your time * Required response. Select one.
dmp
4
0
4
5
0
1
0 I don't know what a data management plan is
1 I don't know if my project has a data management plan
2 My project does not have a data management plan
3 A data management plan was written as part of the bid but has not been updated
4 My project has a data management plan that covers the duration of the research
5 My project has a data management plan that extends beyond completion of the research
Prezi


Question 8: Does the research you are currently working on have a data management plan? * if you are working on more than one project or theme, answer for the one that takes most of your time * Required response. Select one.
dmp
4 0 4 5 0 1
0 I don't know what a data management plan is
1 I don't know if my project has a data management plan
2 My project does not have a data management plan
3 A data management plan was written as part of the bid but has not been updated
4 My project has a data management plan that covers the duration of the research
5 My project has a data management plan that extends beyond completion of the research

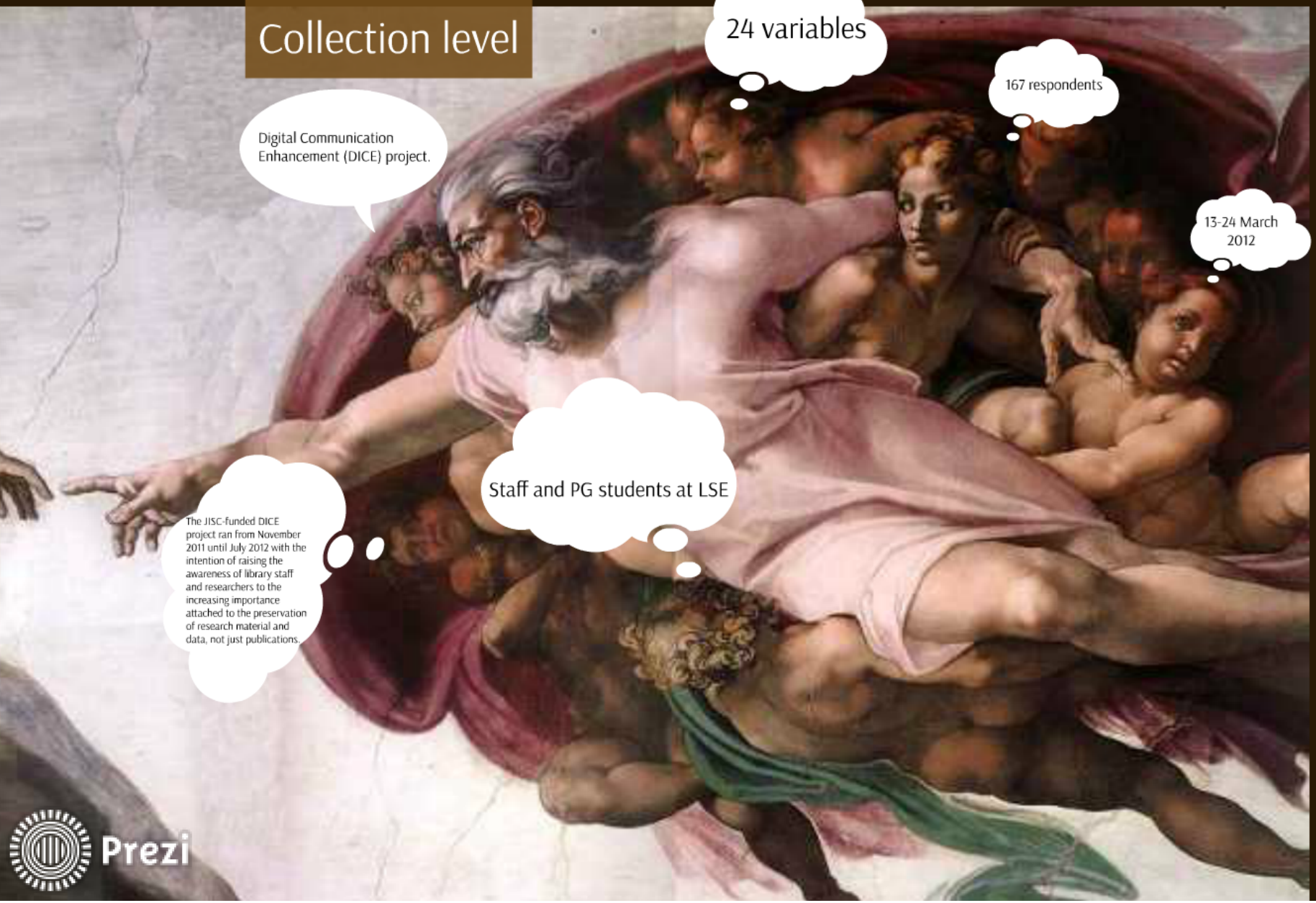

Collection level
24 variables
167 respondents
13-24 March 2012
Digital Communication Enhancement (DICE) project.
Staff and PG students at LSE
The JISC-funded DICE project ran from November 2011 until July 2012 with the intention of raising the awareness of library staff and researchers to the increasing importance attached to the preservation of research material and data, not just publications.
know what a data management
now if my project has a data
ent plan
ect does not have a data
ent plan
nanagement plan was written
the bid but has not been
ect has a data management
covers the duration of the
ect has a data management
extends beyond completion of
rch
Prezi

Collection level

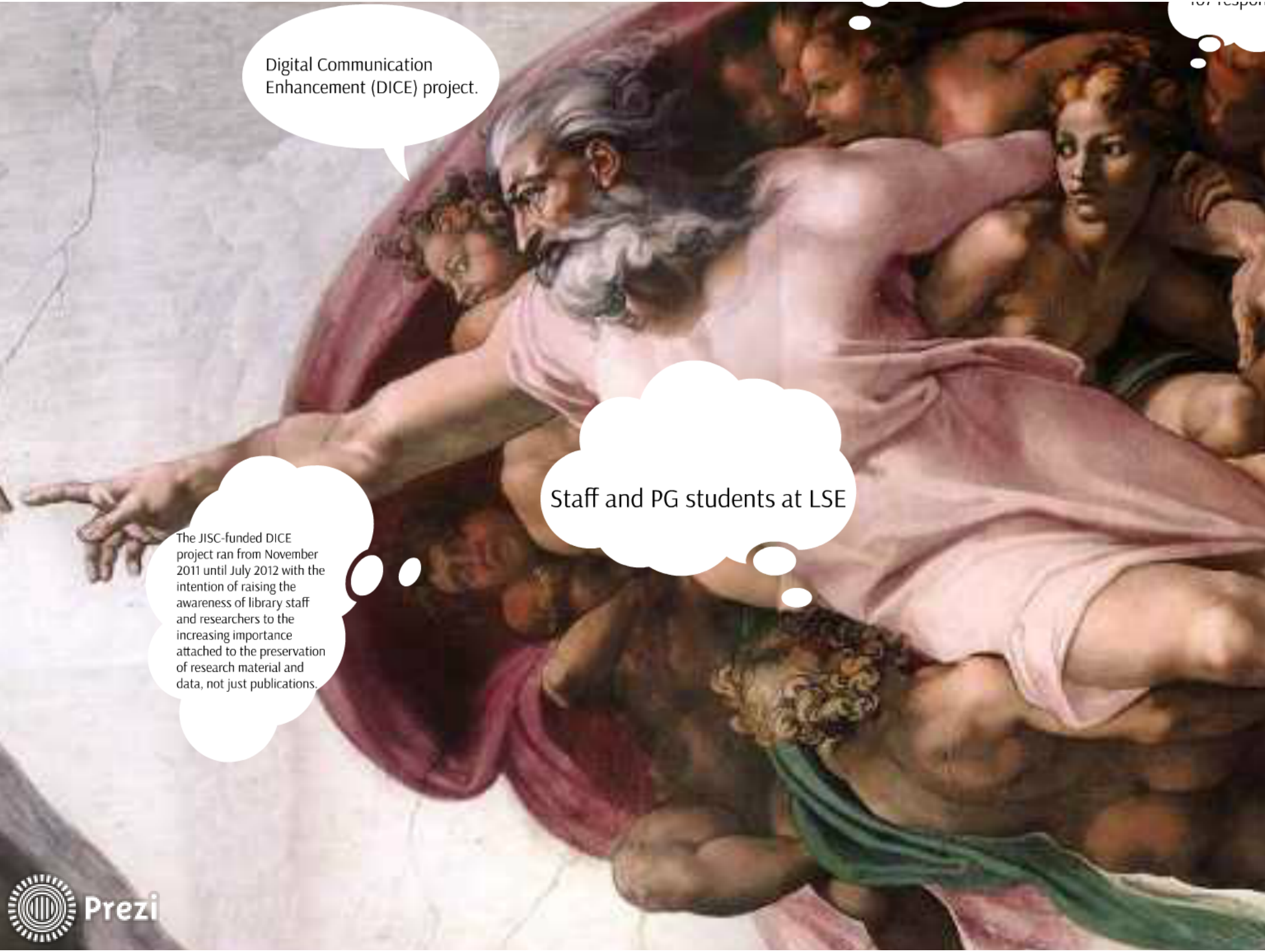


Prezi

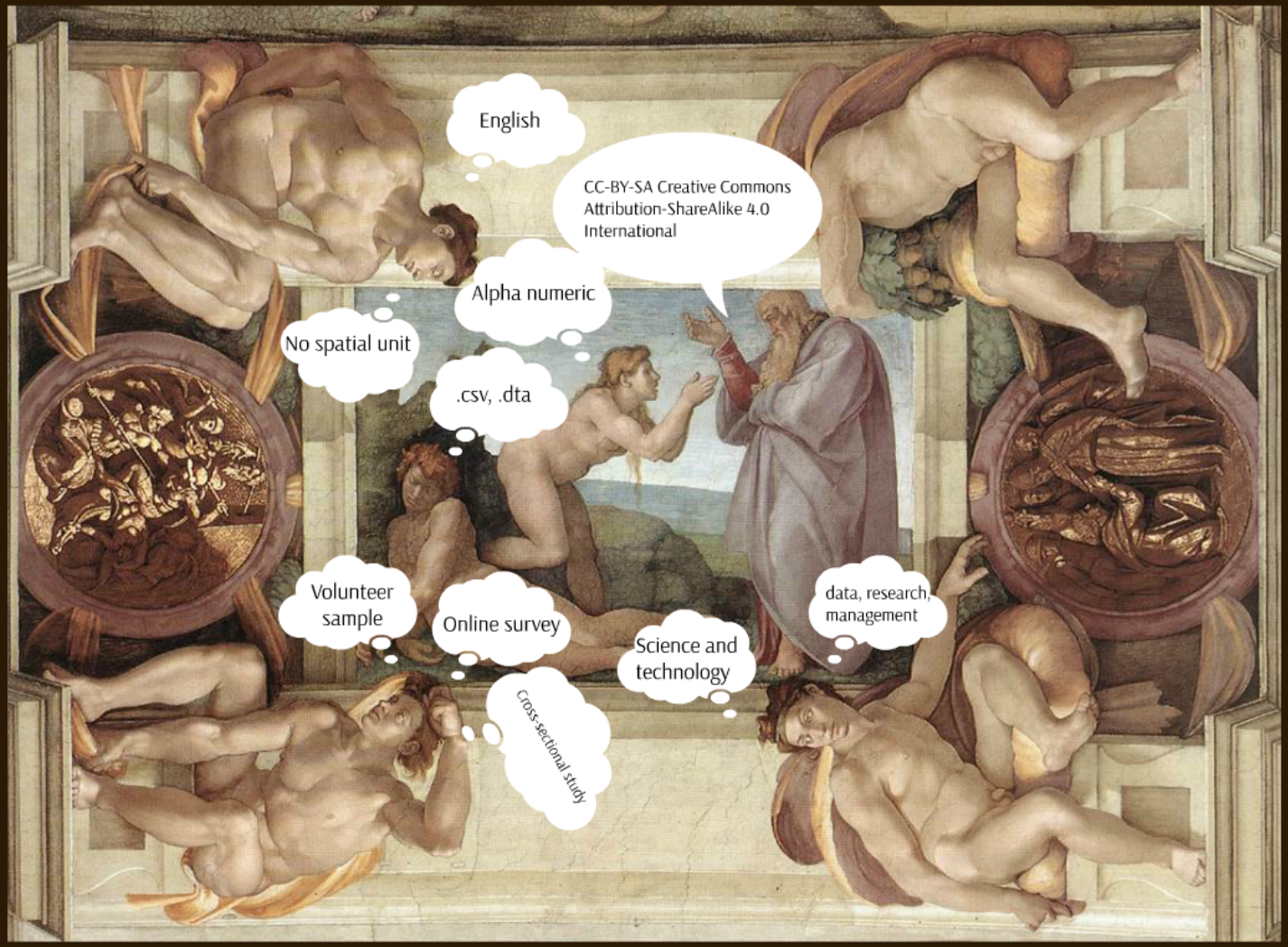
Collection level



Prezi


Digital Communication Enhancement (DICE) project.
Staff and PG students at LSE
The JISC-funded DICE project ran from November 2011 until July 2012 with the intention of raising the awareness of library staff and researchers to the increasing importance attached to the preservation of research material and data, not just publications.
Prezi

"Creación de Adán (Miguel Ángel)" by Michelangelo Buonarroti - See below.. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Creaci%C3%B3n_de_Ad%C3%A1n_(Miguel_%C3%81ngel).jpg#mediaviewer/File:Creaci%C3%B3n_de_Ad%C3%A1n_(Miguel_%C3%81ngel).jpg



"Michelangelo, Creation of Eve 00" by Michelangelo Buonarroti - Web Gallery of Art: Image Info about artwork. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Michelangelo,_Creation_of_Eve_00.jpg#mediaviewer/File:Michelangelo,_Creation_of_Eve_00.jpg

Prezi

What is metadata?

Metadata structure

Types of metadata

Conclusion

Metadata

FOSTER

Discovering Open Practices for Early Career Researchers
4 September 2014

Laurence Horton
Data Librarian, The London School of Economics and Political Science
L.Horton@lse.ac.uk

This presentation is licensed under a Creative Commons Attribution 4.0 International license.

Prezi

"Gustav Klimt 016" by Gustav Klimt - http://www.belvedere.at/jart/prj3/belvedere/main.jsp?userid=&mandator=1&navigation=1&content_id=10000&content_type=1&content_id_parent=10000&showpage=1&kid=1&kunstwerk_id=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent_parent=10000&kunstwerk_id_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent=10000&kunstwerk_id_parent

What information do I need to understand, evaluate, and build upon data without having to ask for additional information from the author?

"Metadata is a love note to the future" by Playing Futures: Applied Nomadology https://www.flickr.com/photos/centralasian/8071729256/ available under CC-BY 2.0 license https://creativecommons.org/licenses/by/2.0/

Are there existing standards I can use to collect or structure that information?

Prezi