

ChIA Image Classification and Annotation

Yalemisew Abgaz¹, Amelie Dorn², Japesh Methuku¹, Renato Rocha
Souza² Gerda Koch³

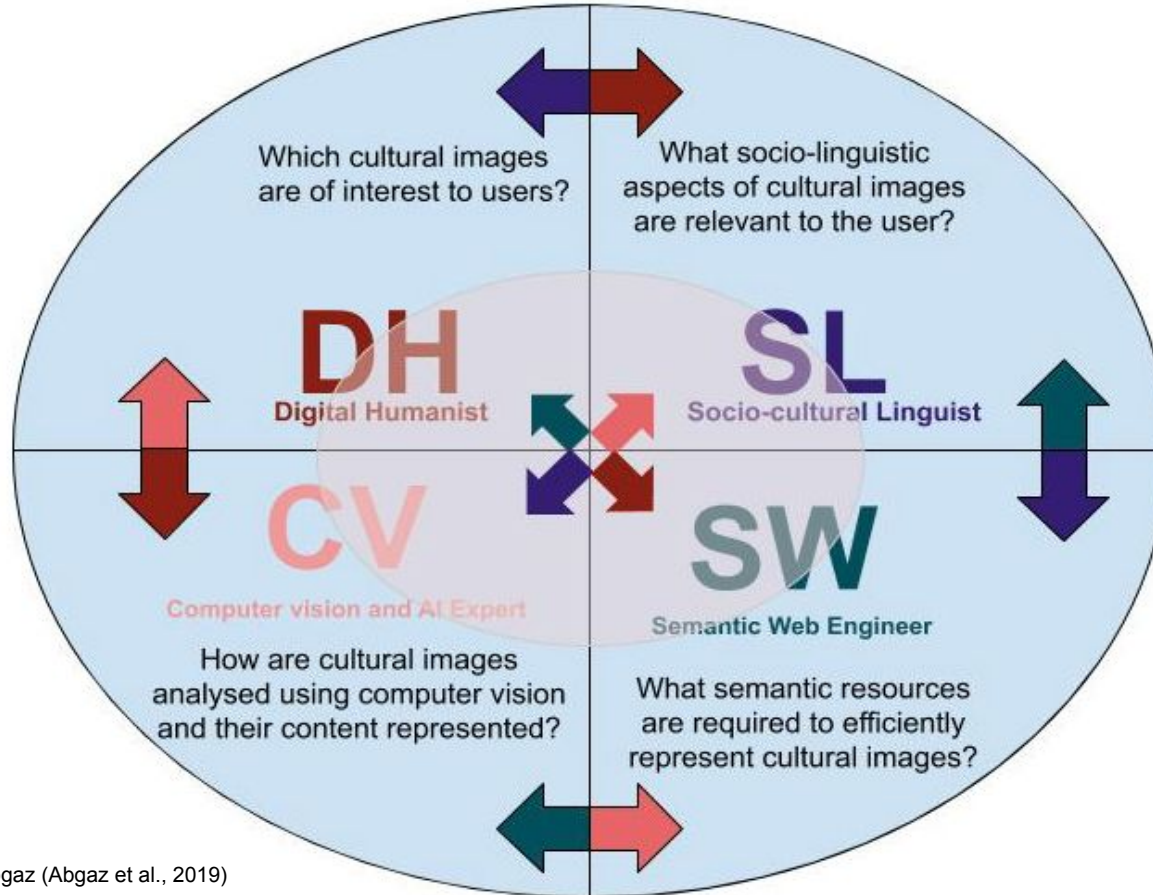
¹Dublin City University, ²ACDH-CH OeAW, ³Europeana Local - AT

Background

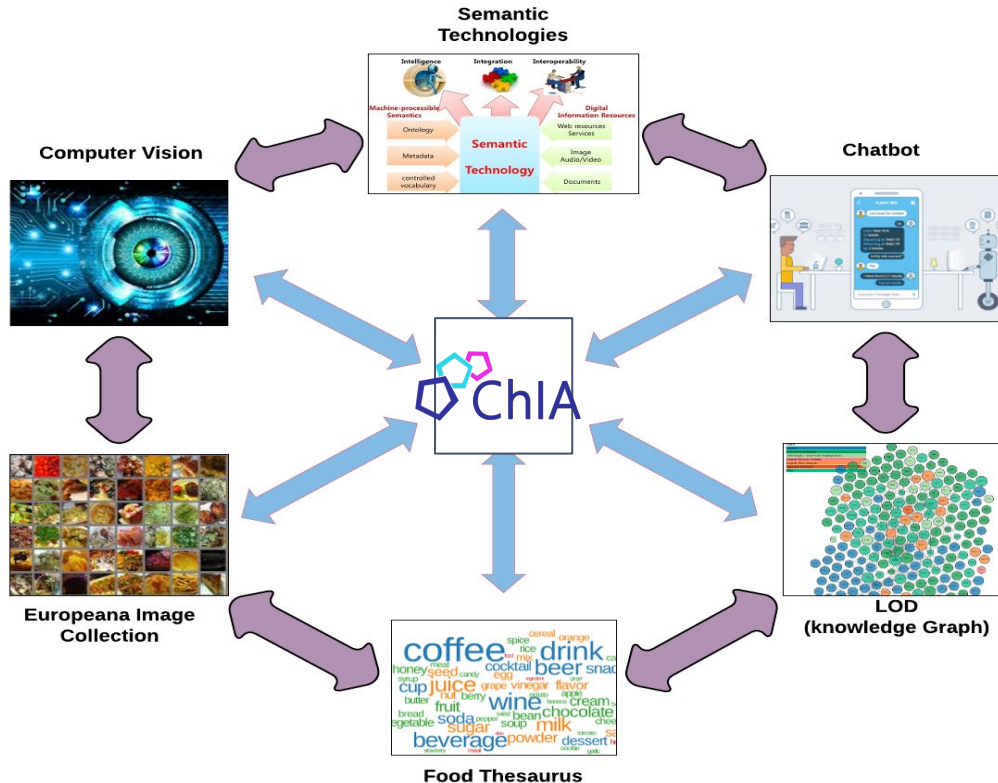
ChIA

- Interdisciplinary Digital Humanities project (2019-2021)
- Involved expertise:
Digital Humanities, AI & NLP (ACDH-CH OeAW, AT)
Semantic technologies (Adapt Centre, IE)
Cultural Image aggregation (Europeana Local - Österreich, AT).
- Project aim & results: **the ChIA system** - enabled increased access and analysis possibilities of cultural (food) images for content providers and educational purposes.

ChIA Collaboration



ChIA Components



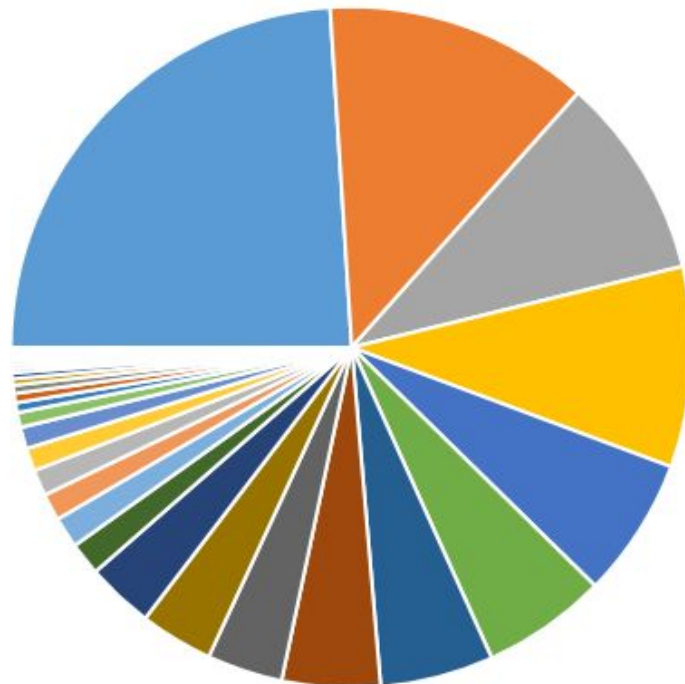
ChIA Data set

Europeana data set

Total: 58.6 Mio digital objects

Includes: 34.2 Mio digital images

from: 3.500 institutions in 42 countries



The Problem

Dataset:

- The dataset is very rich with cultural heritage concepts.
- It represents the social, cultural, economical, political, etc. aspects of a society in time.
- The images appear as photographs, drawings, paintings etc.





Image source: Banquet Still Life (Adriaen Van Utrecht); CC-BY-PD



JUST A LITTLE TRICK....



Image source: <https://www.europeana.eu/de/item/9200579/i83efggh> (Wellcome Collection, CC-BY)

The Problem

The current image collection:

- The collection focuses on bibliographic metadata
- Wide range of cultural information does not get sufficient coverage.

The Problem

Metadata

- In most cases content descriptions use iconographic phrases like “fruits”, “flowers”, “still life”.



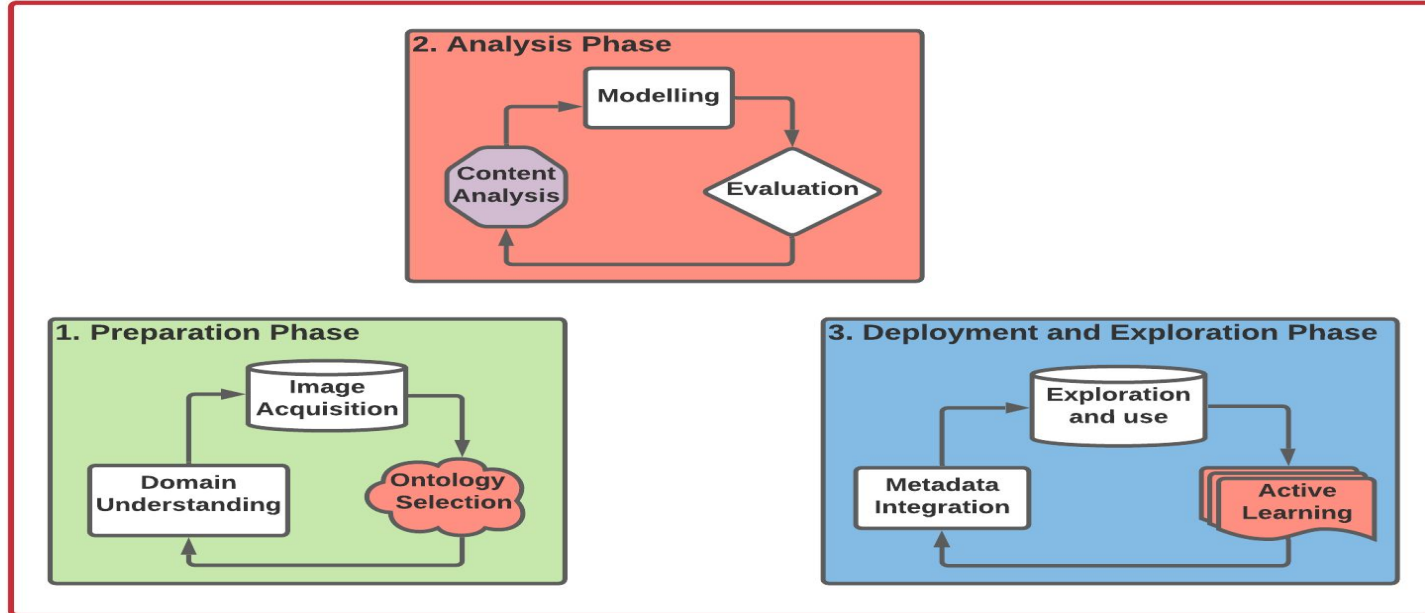
about:	/proxy/provider/2059517 /data_foodanddrink_WAGMU_op725	
dcCreator:	def:	STEELE; E
dcDate:	def:	1889
dcDescription:	def:	Oil painting showing a vase of flowers and a bowl of fruit on a table, more fruit lies on the table.
dcFormat:	def:	image/jpeg

The Problem

Vocabularies

- No specific food or drink related vocabularies are used.
- Irregularity in the use of semantics across the dataset.
- Abstract cultural aspects of the images are hidden.

The Approach

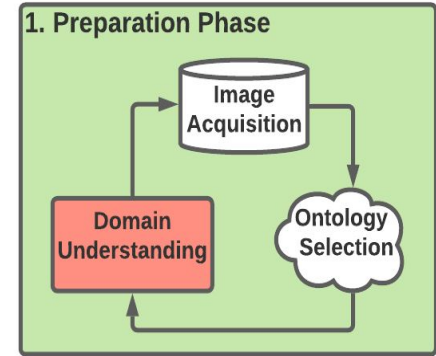


A three-phase-methodology for semantic enrichment of cultural images.

Preparation Phase

Domain Understanding

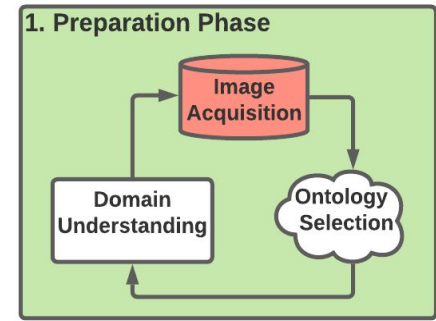
- Understanding the collection.
- Understanding the needs of the users.
- Understanding cultural aspects of the images.



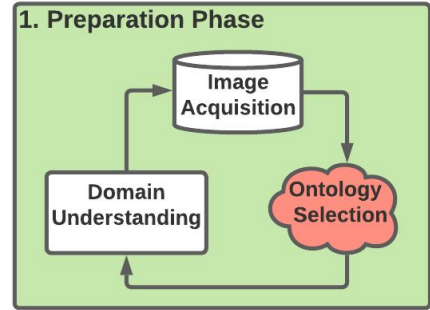
Preparation Phase

Image Acquisition

- Europeana images.
- ChIA search platform.
- Food and drink related images.



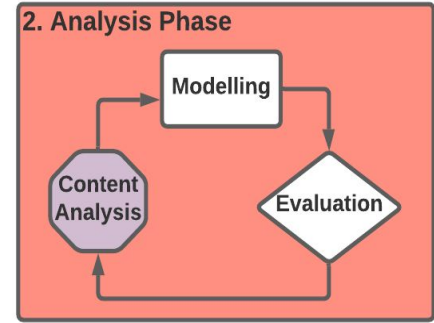
Preparation Phase



Ontology/Vocabulary Selection

- Art and Architecture Thesaurus (AAT).
- Iconclass.
- ChiA food Vocabulary.

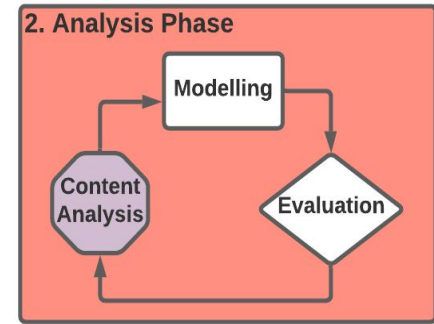
Analysis Phase



Content Analysis

- Cultural aspects
 - Concrete
 - Fruit / Non-fruit related food
 - Fruit: a fruit is something that grows on a tree or bush and which contains seeds or a stone covered by a substance that you can eat. (e.g., strawberry, nut, tomato, peach, banana, green beans, melon, apple).
 - Non-fruit: images that do not feature any type of fruit (for fruit definition see above).

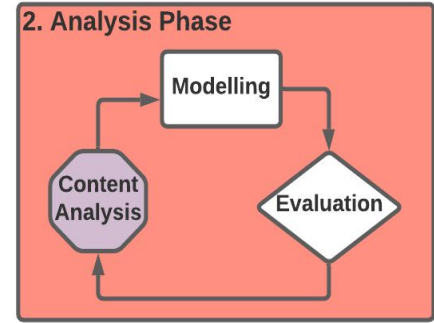
Analysis Phase



Content Analysis

- Cultural aspects
 - Abstract
 - Formal / Informal
 - Formal: arranged in a very controlled way or according to certain rules; an official situation or context.
 - Informal: a relaxed environment, an unofficial situation or context, disorderly arrangement.

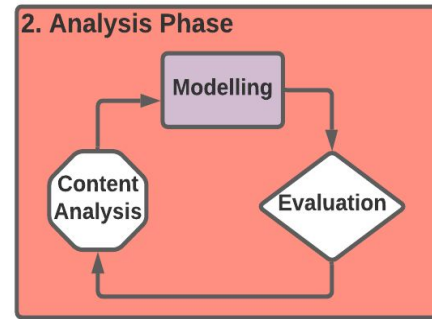
Analysis Phase



Content Analysis

- Cultural aspects
 - Abstract
 - Appealing / Non-appealing
 - Appealing: an image that is a pleasure to look at. A food image that is pleasing to the eye, desirable to eat and good for food.
 - Non-appealing: an image that is not a pleasure to look at.

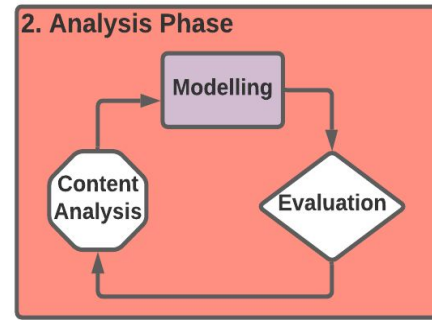
Analysis Phase



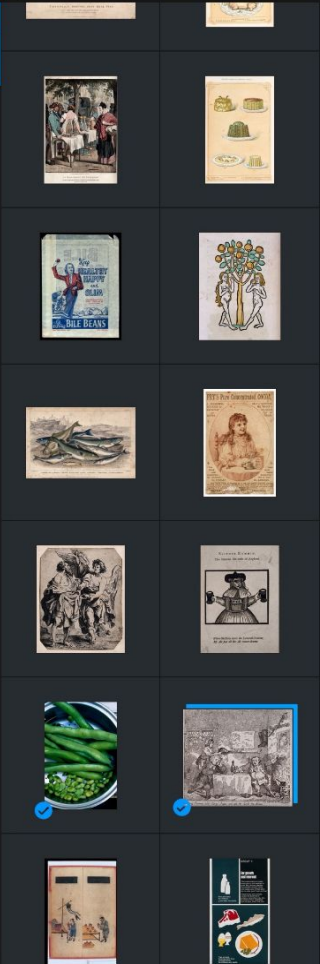
Training Data Annotation

- ~2000 images selected for experiment.
- Four rounds.
- 6 Experts involved.
- 18 Students participated.

Building the training set



- Expert Semantic Annotation
 - **Fruit /Non-fruit** - Highest inter-rater agreement ~92%.
 - **Formal/Informal** - Lowest agreement ~31%.
 - **Appealing/Non-appealing** - Moderate agreement ~53%.





Appealing

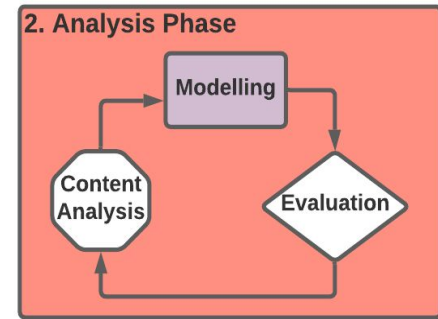
Non-appealing



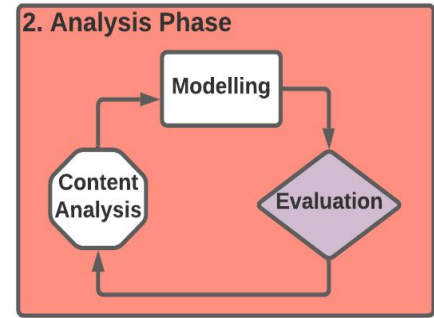
Modelling

Computer Vision training models

- Visual Geometry Group (VGGNet)
- RESNet50
- Inception_v3
- x-ception



Prediction for new images



- **Appealing / Non-Appealing**
 - Prediction by three models.
 - Average prediction is considered.
 - A CSV file for all images along with the prediction is generated.

Model	Training Accuracy	Validation Accuracy	Test Accuracy
Fine tuned ResNet50	83.51%	83.81%	80%
Fine tuned Inception_V3	92.61%	87.62%	90%
Fine tuned Xception	93.2%	88.1%	85.56%

Inception_V3 vs. ResNet50 vs. Xception

Comparison between Fine tuned Inception_v3, ResNet50 and Xception models for images in ChIA framework.

IM



INCEPTION IMAGE CLASSIFIER: OUTPUT

[Appealing]

[Appealing] 92%

[Non-appealing] 8%

Latency: 4.50s

RESNET IMAGE CLASSIFIER: OUTPUT

[Appealing]

[Appealing] 85%

[Non-appealing] 15%

Latency: 2.41s

XCEPTION IMAGE CLASSIFIER: OUTPUT

[Appealing]

[Appealing] 100%

[Non-appealing] 0%

Latency: 2.11s

CLEAR

SUBMIT

gradio

FLAG

Inception_V3 vs. ResNet50 vs. Xception

Comparison between Fine tuned Inception_v3, ResNet50 and Xception models for images in ChIA framework.

IM



INCEPTION IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 74%
[Appealing] 26%

Latency: 0.43s

RESNET IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 69%
[Appealing] 31%

Latency: 0.36s

XCEPTION IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 99%
[Appealing] 1%

Latency: 0.52s

CLEAR

SUBMIT



FLAG

Inception_V3 vs. ResNet50 vs. Xception

Comparison between Fine tuned Inception_v3, ResNet50 and Xception models for images in ChIA framework.

IM



EDIT

INCEPTION IMAGE CLASSIFIER: OUTPUT

[Non-appealing]



Latency: 0.35s

RESNET IMAGE CLASSIFIER: OUTPUT

[Non-appealing]



Latency: 0.27s

XCEPTION IMAGE CLASSIFIER: OUTPUT

[Appealing]



Latency: 0.41s

CLEAR

SUBMIT



FLAG

Inception_V3 vs. ResNet50 vs. Xception

Comparison between Fine tuned Inception_v3, ResNet50 and Xception models for images in ChIA framework.

IM



EDIT

INCEPTION IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 80%
[Appealing] 20%

Latency: 0.35s

RESNET IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 90%
[Appealing] 10%

Latency: 0.28s

XCEPTION IMAGE CLASSIFIER: OUTPUT

[Non-appealing]

[Non-appealing] 98%
[Appealing] 2%

Latency: 0.43s

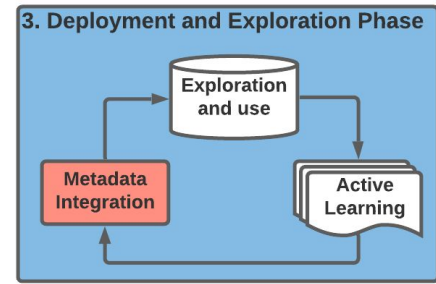
CLEAR

SUBMIT



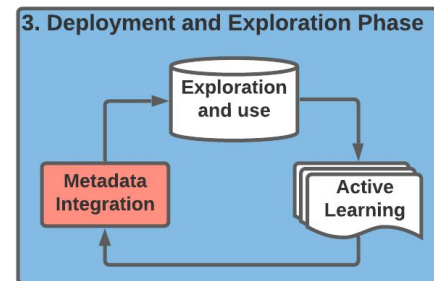
FLAG

Semantic Enrichment



- Generating metadata and Linking images
 - Appealing Images linked to an ontology related to emotion .
 - <http://www.ontobee.org/ontology/MFOEM>
 - Aesthetic Pleasure
 - http://purl.obolibrary.org/obo/MFOEM_000039
- R2RML mapping is employed to the generated CSV files.

Semantic Enrichment



```

<https://www.europeana.eu/en/item/2059511/data_foodanddrink_24255> a
<http://www.europeana.eu/schemas/edm/webResource> ;
<http://www.w3.org/2000/01/rdf-schema#comment> "Appealing" , "Appealing:95.69" ;
<http://purl.org/dc/elements/1.1/description> "Appealing" , "Appealing:95.69" ;
<http://purl.org/dc/elements/1.1/subject> <http://purl.obolibrary.org/obo/MFOEM_000039> .

<https://www.europeana.eu/en/item/2059511/data_foodanddrink_24264> a
<http://www.europeana.eu/schemas/edm/webResource> ;
<http://www.w3.org/2000/01/rdf-schema#comment> "Appealing" , "Appealing:96.33" ;
<http://purl.org/dc/elements/1.1/description> "Appealing" , "Appealing:96.33" ;
<http://purl.org/dc/elements/1.1/subject> <http://purl.obolibrary.org/obo/MFOEM_000039> .

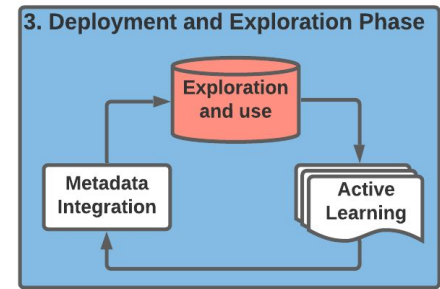
<https://www.europeana.eu/en/item/2059511/data_foodanddrink_24245> a
<http://www.europeana.eu/schemas/edm/webResource> ;
<http://www.w3.org/2000/01/rdf-schema#comment> "Non-appealing" , "Non-appealing:81.24" ;
<http://purl.org/dc/elements/1.1/description> "Non-appealing" , "Non-appealing:81.24" .

<https://www.europeana.eu/en/item/2059511/data_foodanddrink_24263> a
<http://www.europeana.eu/schemas/edm/webResource> ;
<http://www.w3.org/2000/01/rdf-schema#comment> "Appealing" , "Appealing:91.9" ;
<http://purl.org/dc/elements/1.1/description> "Appealing" , "Appealing:91.9" ;
<http://purl.org/dc/elements/1.1/subject> <http://purl.obolibrary.org/obo/MFOEM_000039> .
  
```

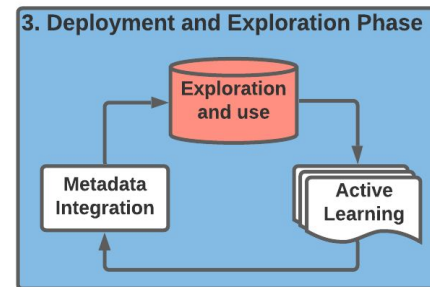

Semantic Exploration

- Used to expand the image search
 - SPARQL-queries

```
prefix obo: <http://purl.obolibrary.org/obo/>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix dc: <http://purl.org/dc/elements/1.1/>
prefix edm: <http://www.europeana.eu/schemas/edm/>
select ?subject ?predicate ?object
where{
?subject ?predicate ?object.
?subject rdf:type edm:webResource.
?subject dc:subject obo:MFOEM_000039.
}
limit 15
```



Semantic Exploration



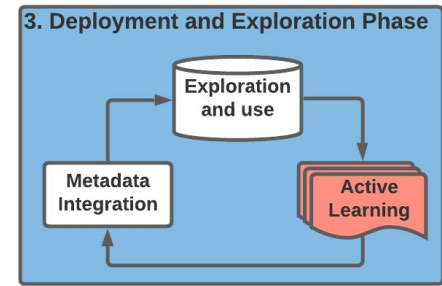
- Used to expand the image search
 - SPARQL-queries - group appealing images together

```

<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> dc:subject obo:MFOEM_000039
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> dc:description Appealing
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> dc:description Appealing:95.69
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> rdfs:comment Appealing
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> rdfs:comment Appealing:95.69
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24255> rdf:type edm:webResource
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> dc:subject obo:MFOEM_000039
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> dc:description Appealing
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> dc:description Appealing:96.33
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> rdfs:comment Appealing
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> rdfs:comment Appealing:96.33
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24264> rdf:type edm:webResource
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24263> dc:subject obo:MFOEM_000039
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24263> dc:description Appealing
<https://www.europeana.eu/en/item/2059511/data\_foodanddrink\_24263> dc:description Appealing:91.9
  
```

Active Learning

- Continuous development
- Continuous feedback
 - We continuously collect data about what users think about the food images.
 - We are considering additional cultural features to enrich the semantics.



Observations

- Lessons learned
 - Annotator agreement.
 - Content Analysis is a key to deliver what users need.
 - Building a sound training dataset is key.
 - CV results are as good as human annotation.
 - Incorporate active learning

Observations

- Continuous integration
 - Expert annotation.
 - Semantic enrichment.
 - user-defined explorations.
 - Chatbot support.

The Future

- Expanding the CV and image classification application to other concepts.
- Further collection of training set.
- Semantic interlinking.
- Supporting exploration.

Thank You

Check out our website

<https://chia.acdh.oeaw.ac.at>

Our publication of the work

<https://doi.org/10.3390/jimaging7080121>