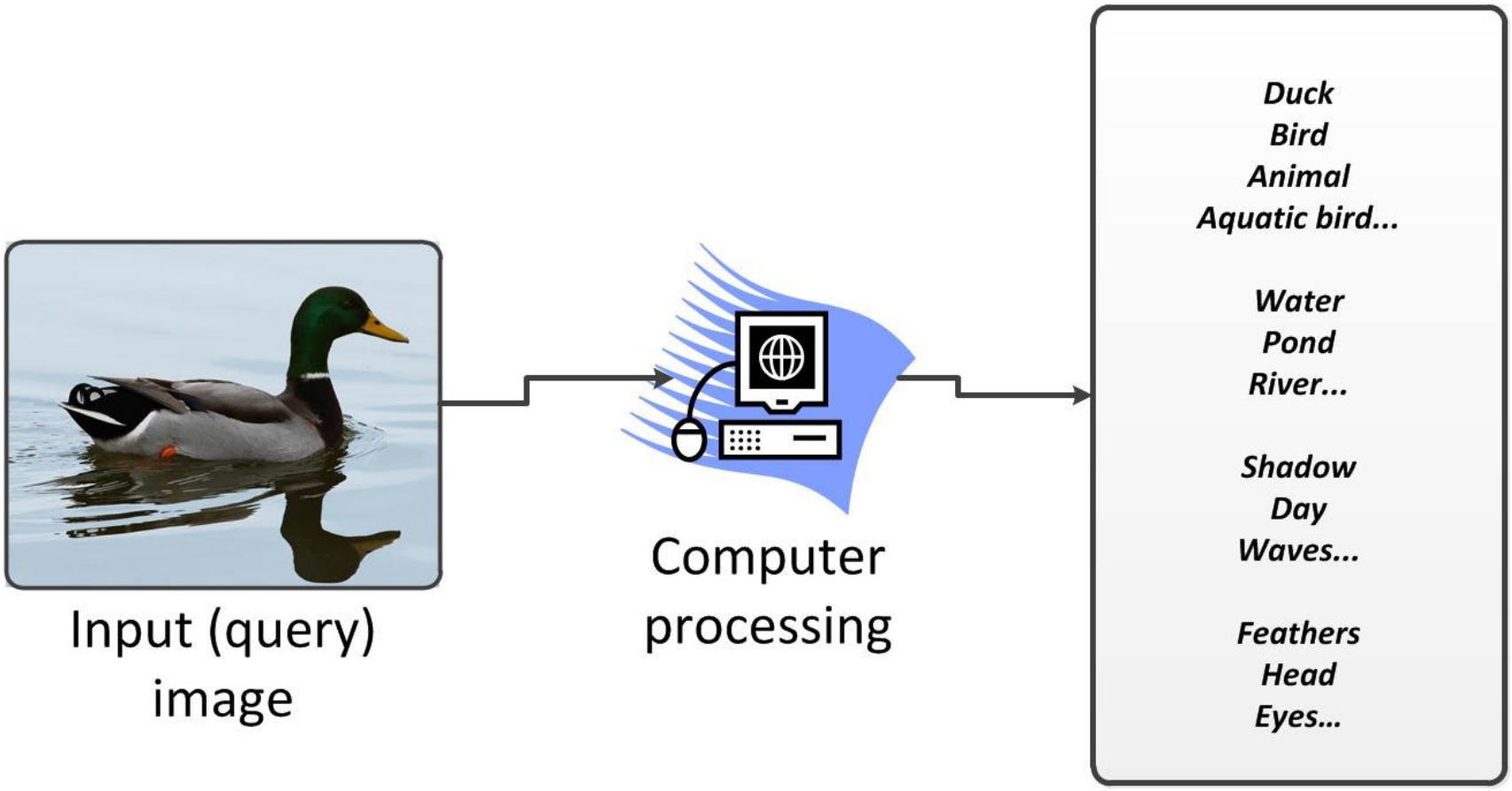
Annotation Framework – Recent state of the application

Honza Botorek, Petra Budíková, Michal Batko

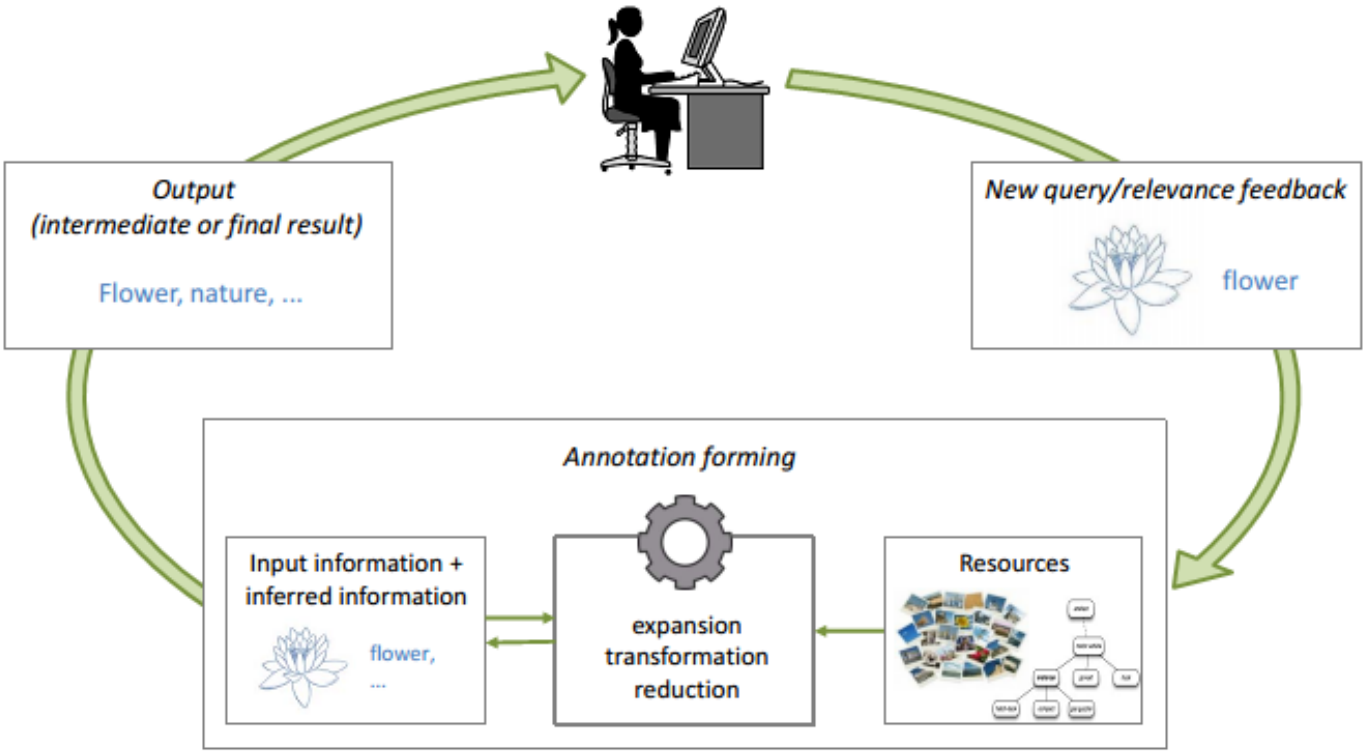
Automatic image annotation – What is it?



- Main goal = annotate unknown images with relevant descriptive words

Annotation Framework: Current Approach I.

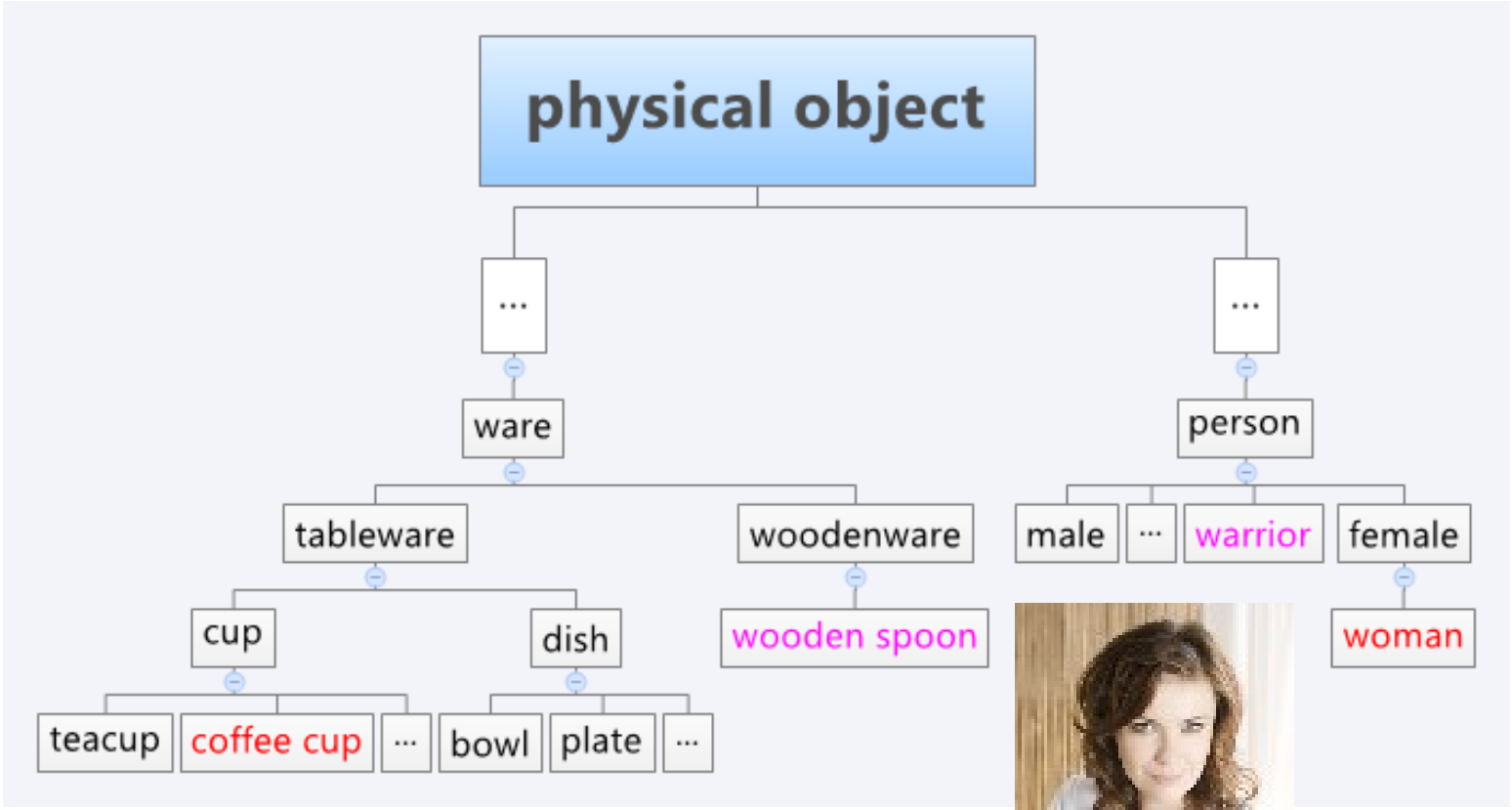
- Given image – CBIR performed (MUFIN) – textual data are retrieved (descriptions)
  - iteratively processed by expansion/transformation/reduction members => output
  - Us



Annotation Framework: Current Approach II. Annotation Forming

- Fundamental technology: **Wordnet**
- Words are interrelated by meanings – basic relation = synonymy
  - synonymous set of words - synset (car, auto, automobile, machine = 1 object)
- Important relations utilized to group words together:
  - Hypernymy
    - Dog IS-A Animal
  - Hyponymy
    - Animal HAS-DESCENDANT dog
  - Meronymy
    - Dog HAS-PART tail, head, ears…
  - „Gloss relation“
    - „… (dog) has been domesticated by man since prehistoric times … „
      - => *domesticated, man, prehistoric, times*
  - When relation between 2 words is found, group is formed = 2 words are related

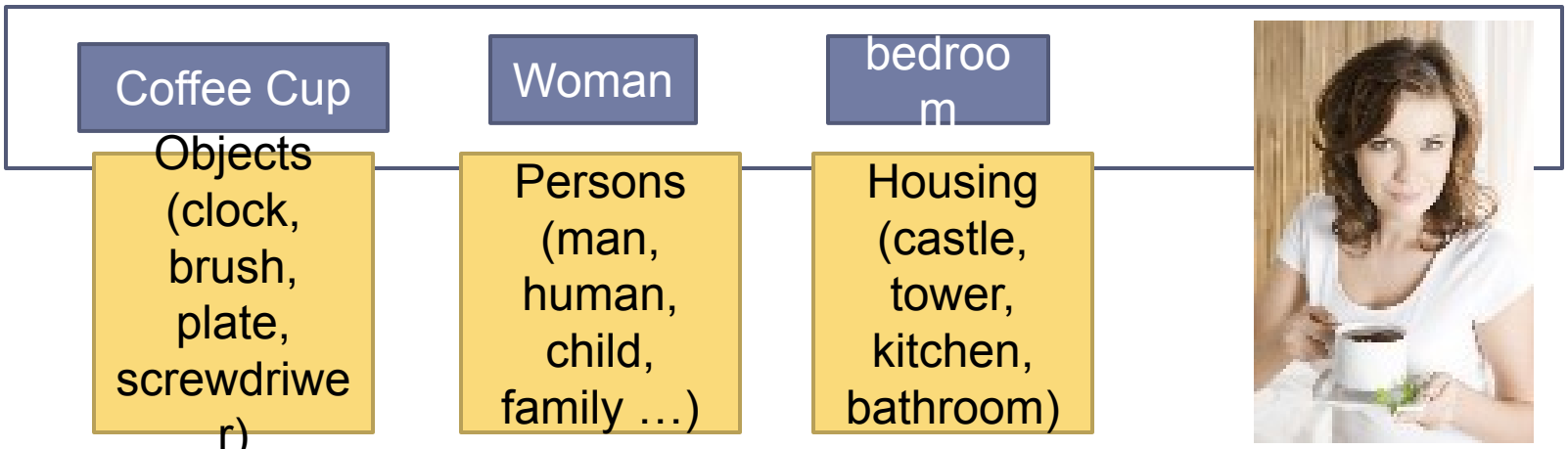
WordNet hypernymy tree – example



… coffee cup, wooden spoon, warrior, woman …

Limitations of current solution

- Grouping forms large set of words – mutually unrelated



- Not structured output from the framework
  - Currently : (dog, puppy, boy, son, child, house)
  - Idea: (animals:{dog, puppy}, persons:{boy, child, son}, buildings{house}
- Accuracy of annotation is not very high
- Annotation Forming tools – space for improvement

Proposed solution I.

- Define a hierarchy of categories that enables to refine annotation results
  - 2 phases: select proper categories; use categories to enrich original query
  - Easier and more accurate annotation process
    - Structured output
  - Ground truth for testing

- User-driven relevance feedback
  - Idea: Iterative process of image annotation
  - Solid hierarchy background is needed

Proposed solution II.

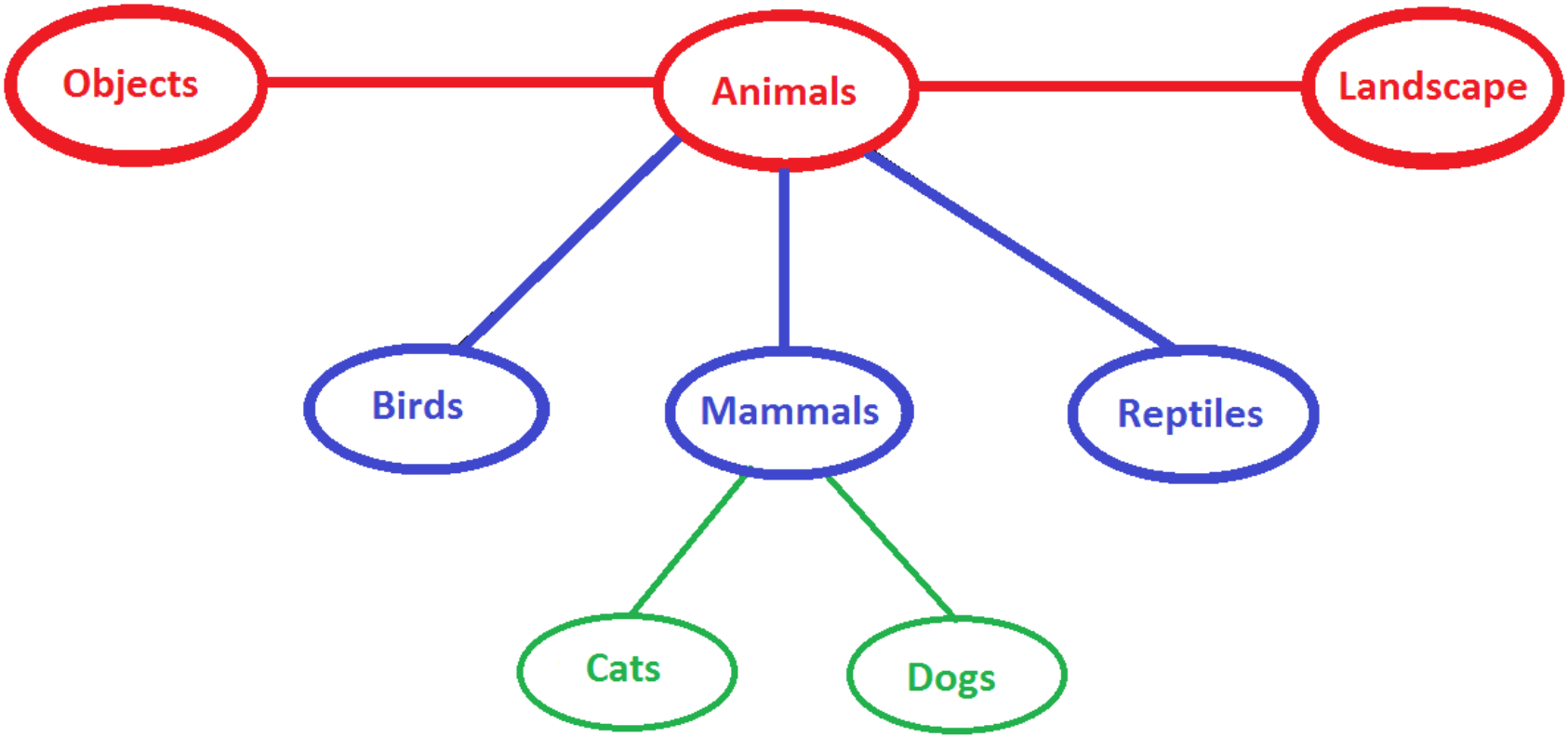
- Add other sources of information (relations among words/objects)
  - Wikipedie: project DBPedia
    - Final thesis topic
- Extend classifiers utilization
  - Indoor x outdoor; buildings detection…
  - OpenCV: Good support for classifiers developing
    - Final thesis topic?

Category tree challenges

- How to create/select ontology categories?
- How to use such categories in the annotation process?
- Which relations encode into ontology?

Categorization – Ontology Motivation

- “Map words into categories to improve a quality of image annotation”



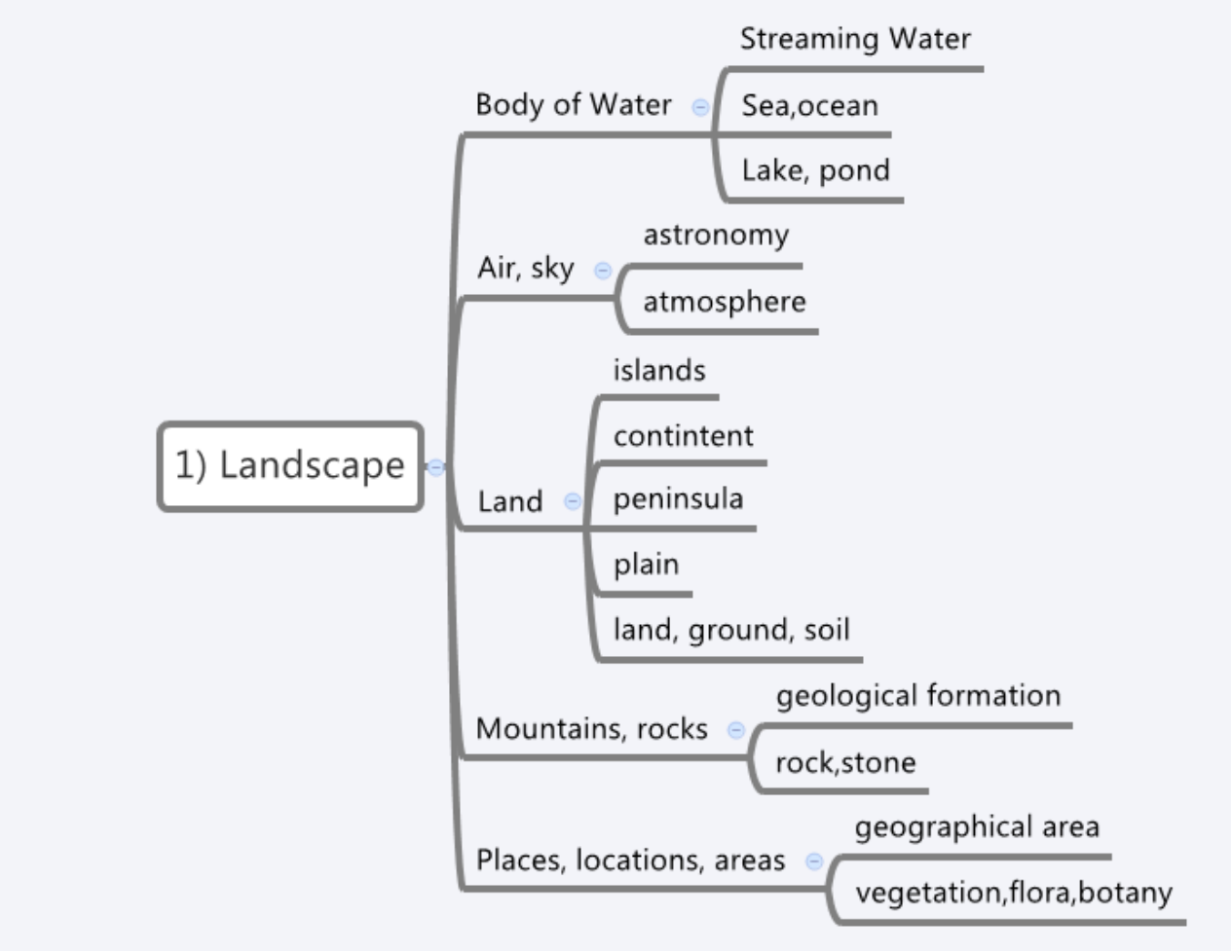
What is an ontology

- *„An ontology is a set of concepts – things, events and relations. These concepts form a vocabulary for exchanging information.“*
- Relations encoding:
  - **<fruit> <subclass_of> <food>**
  - <Movie XY> <hasStar> <John Newman>
    - < hasStar > <domain> <movie>
    - < hasStar > <range> <actor>
- No general ontology exists
- ImageCLEF, LSCOM, DBpedia ontology
- Some examples of specialized ones
  - Food, family, wines, financial institutions…

How to create an ontology? Category Tree I

- Map categories to vital synsets in WordNet structure
- Fundamental/root categories
  - 13 selected (animals, objects, landscape…)
  - sub-categories for each „root category“
    - Animals – birds, mammals, reptiles
      - Mammals – cats, dogs…
  - How categories were selected?
    - Wordnet – parsing of noun synsets with a high number of hyponyms
    - Large ontologies checking (LSCOM, ImageCLEF)

Category Tree II – part of the tree



What relations incorporate into ontology? Category tree III

- IS-A relation: Fundamental requirement to hold relations of the type:
  - Fiats ARE cars; Cars ARE Vehicles; Vehicles ARE Objects …
  - From more exact categories to more general ones
- Incorporation of foreign ontologies
  - More specialized hiearchies to some narrow field (eg. food, cars)
- Relations encoding into the tree
  - opposites (black vs white)
  - „person EAT food“ etc.

How to use the ontology? Category Tree IV

Query image

Produce Annotation

Woman, Cup, Coffee cup, Morning, House, Home, Wife, Bedroom, Table, Sky, Eye contact, Hair, Building, Travelling, Relationship, Person, Man, Indoor, Smile, Drinking, Wellness, Girl, Beauty, BEd, Relax, Sandwiches, Mug

OBJECTS
p = 0,6

CONTAINERS (0,8): Coffee cup, mug, cup

FURNISHING: Bed, table

PERSON, SOCIETY
p = 0,8

SINGLE PERSON: Woman, girl, wife, person

CONSTRUCTION
p = 0,05

BUILDINGS: Building, home, indoor, bedroom

Annotation Framework

Query image

Similar images

MUFIN SIMILAR IMAGE SEARCH

Query image

Person, object

X

woman, coffee cup

MUFIN SIMILAR IMAGE SEARCH

Similar images

Query object

Annotation Framework

Produce annotation

Woman, Cup, Coffee cup, mug, person, cofee, tea, girl, dishware

Produce new Query object

Summary

- Ontology constructed on WordNet structure is designed
- The ontology helps us to improve annotation results
  - It can produce more general or more specific annotation
  - Different kinds of relations can be encoded
- The ontology is extensible and customizable

- Near future work
  - Implement the ontology into the annotation process
  - Incorporate another ontologies

- Future work
  - Employ the ontology for user relevance feedback