


# Semantic search-based image annotation



Petra Budíková, FI MU  
CEMI meeting, Plzeň, 16. 4. 2014

# The annotation problem

- Formalization

- The annotation problem is defined by a **query image  $I$**  and a **vocabulary  $V$**  of candidate concepts
- The **annotation function  $f_A$**  assigns to each concept  $c \in V$  its probability of being relevant for  $I$

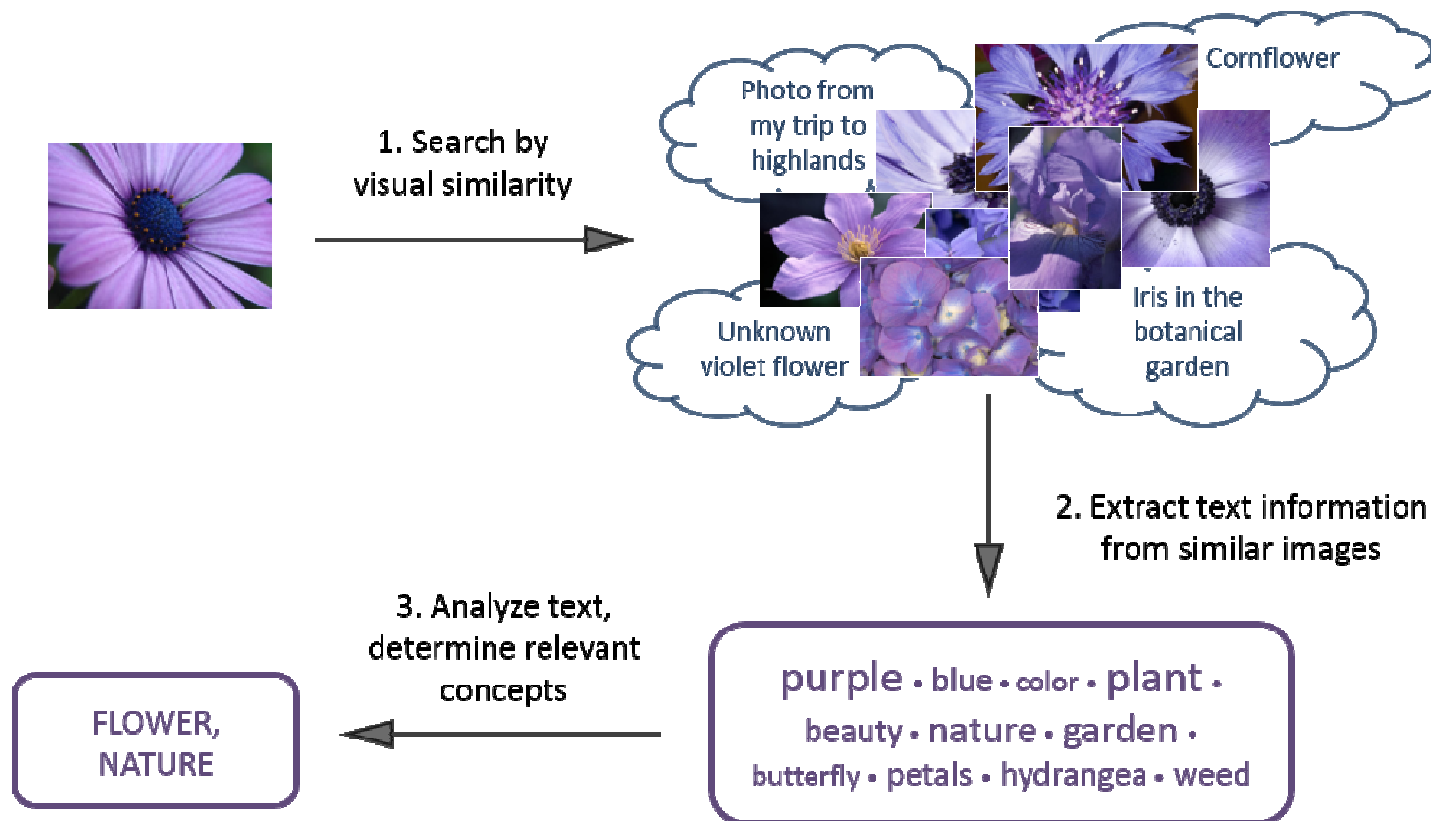


$V = \{\text{flower, animal, person, building}\}$

- Basic possible approaches

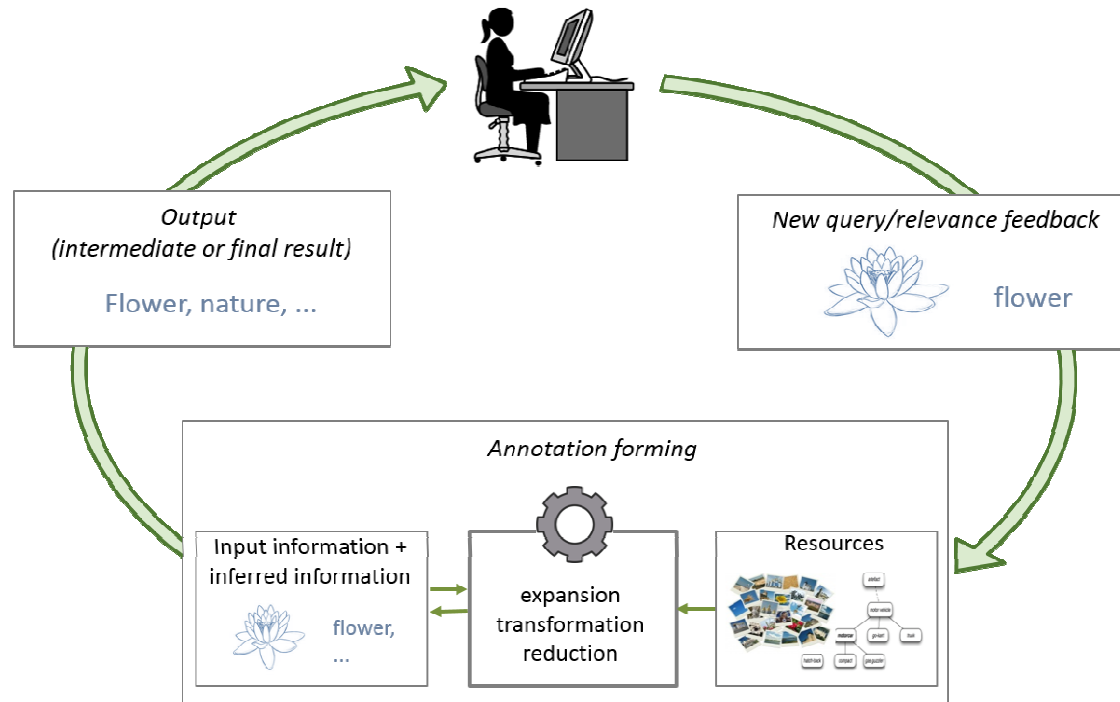
- Model-based annotation**
  - Train classifiers
  - Suitable for tasks with smaller dictionaries and available training images (e.g. medical image classification)
- Search-based annotation**
  - Exploit results of similarity search in annotated images
  - Suitable for tasks with wide dictionaries (e.g. image annotation for web search)

# Search-based annotation in a nutshell



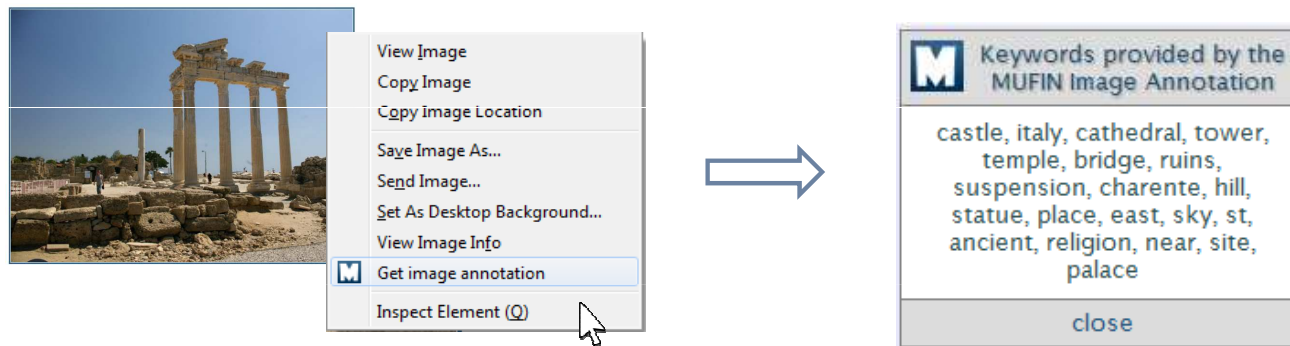
# Our vision

- Next generation of similarity-based annotation
  - Similarity searching
  - Text cleaning
  - Semantic information extraction
  - Classifiers
  - Relevance feedback



# MUFIN Image Annotation

- Already done (paper IDEAS 2013):
  - Modular framework for annotation processing
  - Implementation of basic modules
    - Similarity search, text cleaning, basic WordNet-based semantic processing
  - Working system for keyword annotation with 50-60 % precision
    - Vocabulary V = all English words



- Problems
  - Not precise enough
  - Results too unstructured for practical use
  - Difficult to evaluate

# Current focus

---

- **Hierarchical approach**
  - Vocabulary hierarchically organized
  - WordNet hypernymy/hyponymy tree, ontology
- Semantics-aware processing of similar images' descriptions
  - Study and exploit suitable resources of semantic information
  - Determine the relevance of candidate concepts with respect to semantic relationships
- **ImageCLEF evaluation**
  - ImageCLEF2014: scalability-oriented, no manually labelled training data
  - 100 test concepts, provided with links to WordNet synsets

# ConceptRank

---

- Inspiration: PageRank

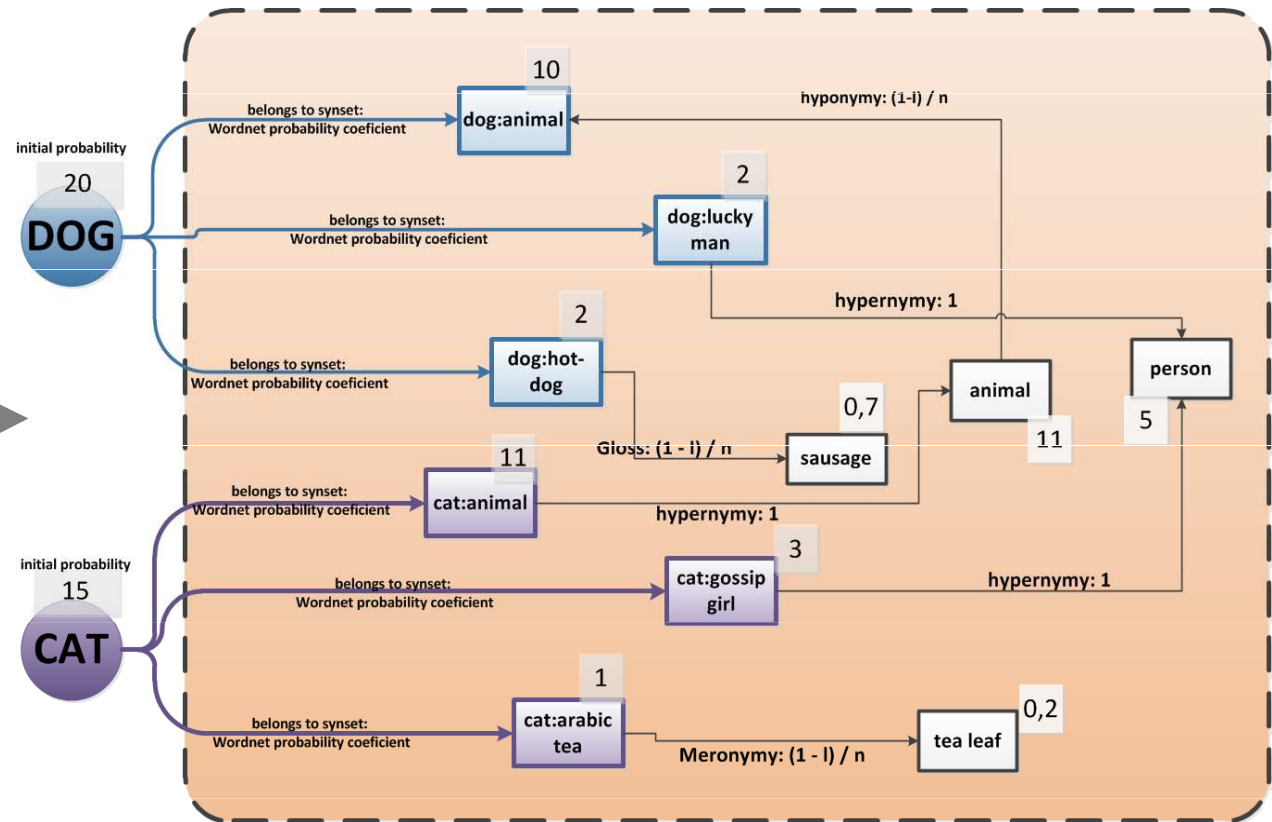
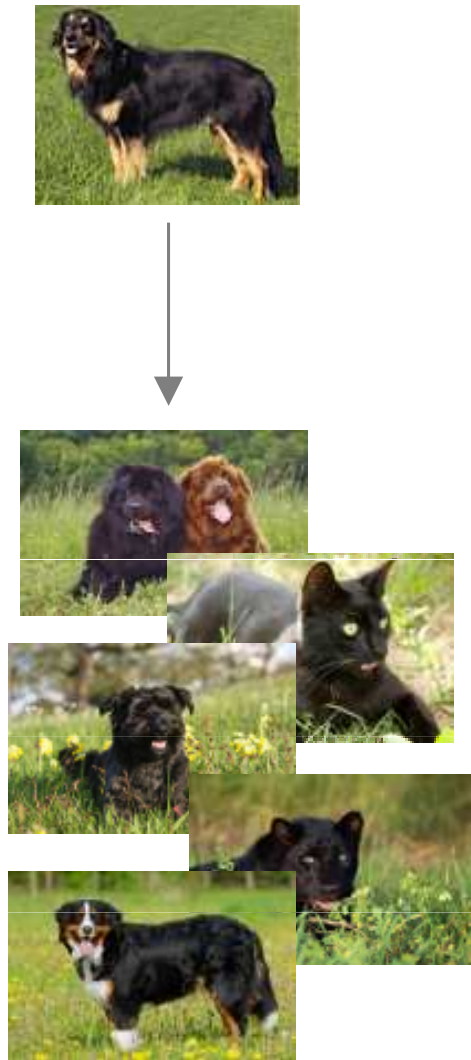
- Importance of a page is derived from the importance of pages that link to it
- Linear iterated process, modelled as a Markov system
- Random restarts to avoid “rank sinks”

$$PR(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{PR(V_j)}{|Out(V_j)|}$$

- ConceptRank idea: Semantic ranking of WordNet synsets

- A Markov system, nodes are formed by WordNet synsets
- Links between nodes connected by some WordNet relationship
  - Weighted according to the type of the relationship
- Random restarts are not weighted uniformly, but reflect the initial weights of synsets as determined by similarity searching

# ConceptRank illustration





# ConceptRank Resources

---

- Content-based image retrieval
  - powered by MUFIN
  - 20M Profiset collection, 250K ImageCLEF training data
- WordNet
  - Standard relationships (hypernymy, antonymy, part-whole, gloss overlap, ...)
  - Word similarity metrics defined on top of hyponymy/hypernymy tree
  - the “language” point of view
- Visual Concept Ontology (VCO)
  - Semantic hierarchy of most common visual concepts, linked to WordNet
  - VCO sub-trees are used to limit the search for WordNet relationships
- Co-occurrence lists for keywords from Profimedia dataset
  - Constructed from very large text corpus (linguists from MFF UK)
    - Corpus size approximately 1 billion words
  - “human/database” point of view

# Cooperation with other CEMI teams

---

- UFAL
  - Information about keyword co-occurrence in text corpora
    - Already part of MUFIN Image Annotation processing
  - Other semantic resources: WikiNet
    - Being studied at UFAL
- ČVUT
  - High-precision classifier for 1000 ImageNet concepts
    - Todo: compare performance of this classifier and MUFIN search-based solution; if complementary, try to combine
  - Image similarity measure derived from the classifier
    - Todo: compare it to MPEG7 similarity utilized by MUFIN Image Annotation

# Questions, comments?

---

