

# Communicative Images

Radek Ošlejšek

`oslejsek@fi.muni.cz`

Faculty of Informatics  
Masaryk University, Brno  
Czech Republic

- ◆ **Motivation:** Looking at a photo from a holiday ten years ago: It is apparently somewhere in mountains, but which mountains and what place? What is that tower in the background? ...
- ◆ **Communicative image** = an “intelligent” image that is able to discuss its content with the user in natural language.



► **Fig.** Photo from a holiday ten years ago

**S:** *This is a photo of a group of your friends standing before the Pernštejn castle. Now, you can open and view the photo or start the communication with image.*

**U:** What is the resolution of this picture?

**S:** *The resolution of this picture is 8 MPix. Enter your next query, please.*

**U:** Who is in the picture?

**S:** *There is a group of your friends. George, John and Elisabeth. There are also four other groups of tourists.*

**U:** Give me the detailed information about this castle.

**S:** *Pernštejn Castle is a castle located on a rock above the village of Nedvedice and the rivers Svratka ...*

**U:** Thank you, bye.

**S:** *Good bye!*



## Requirements:

- ◆ Well defined semantics suitable for automated processing
- ◆ Integration of semantics with graphical data
  - i.e. a graphical format supporting semantics
  - EXIF, SVG, ...
- ◆ Interactive communication by means of natural language
  - written questions/answers, i.e. no speech recognition, no voice synthesis.
  - Information filtering
- ◆ Web environment, social networks, ...
- ◆ Image recognition techniques

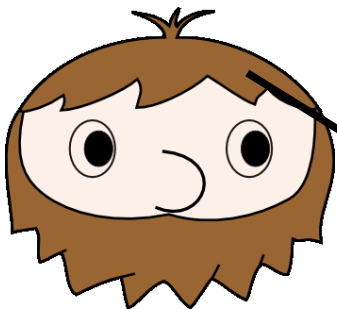
# Ontology-based annotation

## ◆ Unstructured annotation

- Textual description, keywords, etc.
- Adequate for some tasks, e.g. full-text search for relevant images from huge collection
- Insufficient for dialogue-based image investigation

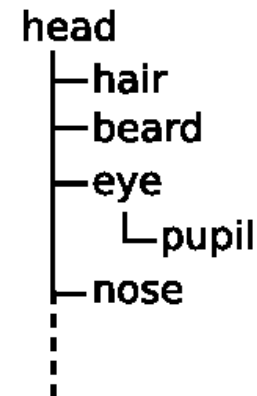
## ◆ Ontology-based structured annotation

- Ontology defines semantics of real object
- An image classifies concrete graphical elements in the ontology

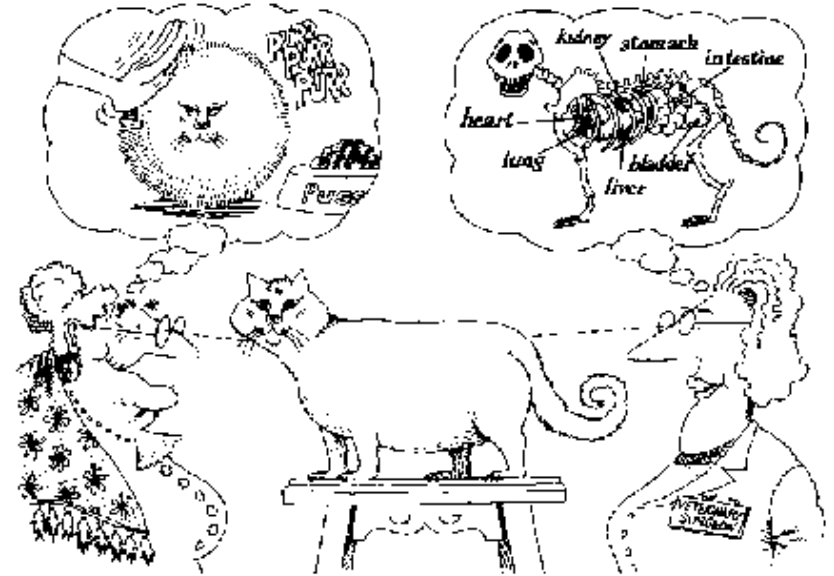


```
<Hair rdf:ID="head">
  <hasColor rdf:resource="#brown" />
</Hair>
```







### Ontology fragment



- ◆ Classes, properties and individuals.
- ◆ Shared knowledge stored in the ontology vs. annotation data stored in the image
- ◆ Problem of abstraction: dangerousness vs. species

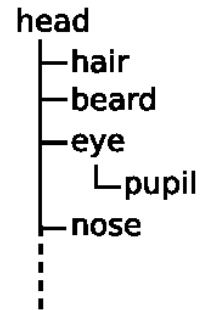


- ◆ Problem of granularity and accuracy of semantic data
  - an Object with description "Boeing 747 of Korean airlines that carried us to Seoul",
  - an Airplane with type set to "Boeing 747" and description "Airplane of Korean airlines that carried us to Seoul",
  - an Airplane with type set to "Boeing 747", airlines set to "Korean" and description "The airplane that carried us to Seoul"...

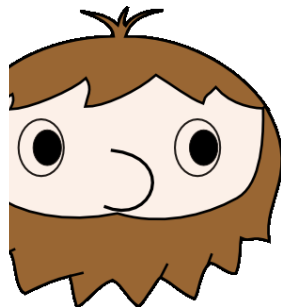
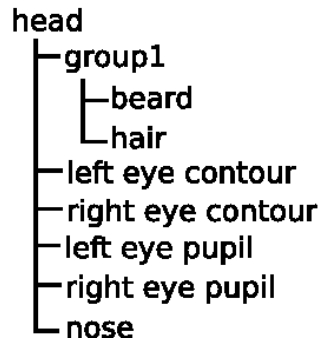
-  OWL brings mathematical formalism with automatic inference
-  Structured knowledge prevents chaos in terminology
-  Shared multilingual knowledge
-  Choice of suitable abstraction of the ontology
-  Building and extending the ontology
-  Laborious annotation process

# SVG and OWL Integration

## Ontology fragment



## Annotated scene graph



## SVG fragment:

```

<g id="head">
  <ellipse ...head geometry definition... />
  <g id="lefteye">
    <ellipse ...eye contour definition without specific id... />
    <ellipse id="leftpupil" ...pupil definition... />
  </g>
  ... scene graph definition continues here ...
</g>
  
```

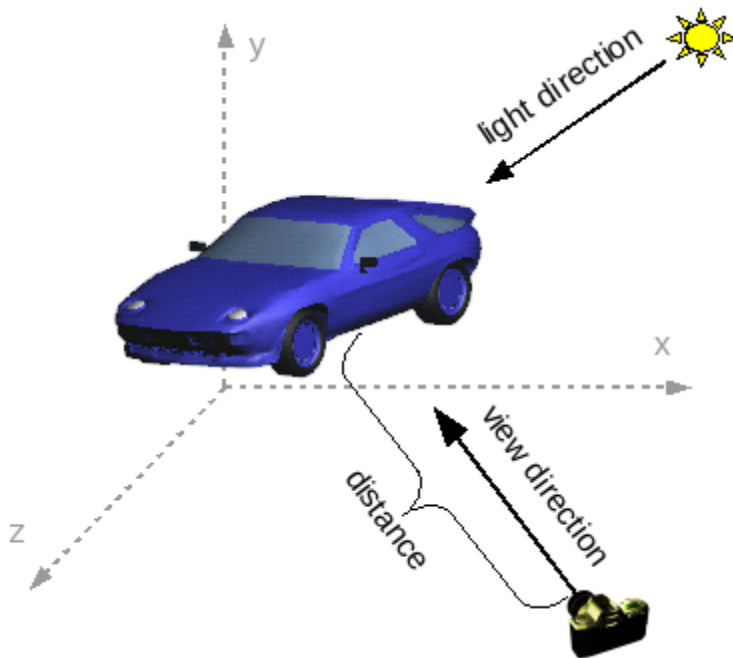
```

<metadata id="ANNOTATION_METADATA">
  <rdf:RDF>
    <owl:Ontology rdf:about="">
      <owl:imports rdf:resource="http://owl.com/ontology.owl" />
    </owl:Ontology>

    <Head rdf:ID="head" />
    <Eye rdf:ID="lefteye" />
    <Pupil rdf:ID="leftpupil" />
    ... classification continues here ...
  </rdf:RDF>
</metadata>
  
```



- ◆ Handles common visual characteristics.
- ◆ Prescribed properties are based on the principles of 3D image synthesis.



Class hierarchy    Class hierarchy (inferred)

Class hierarchy: 'noticeable size'

- Thing
  - color ≡ 'dominant color'
  - 'depicted object'
  - depth
  - image
  - 'object at location'
  - sector
  - 'visual feature'
    - 'dominant color' ≡ color
    - genre
    - 'noticeable shape'
    - 'noticeable size'
    - 'shading effect'
      - 'ambient light'
      - 'light direction'
    - 'view direction'

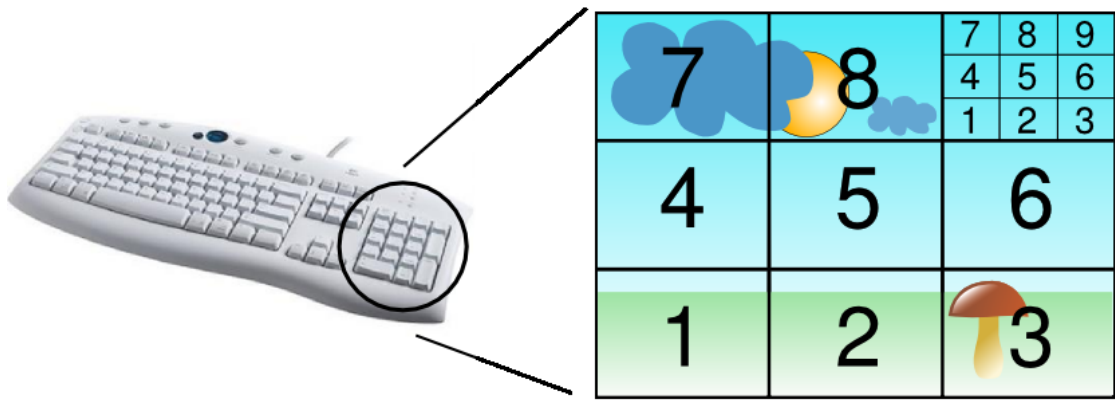
Members list    Members list (inferred)

Members list:

- ◆ big
- ◆ giant
- ◆ high
- ◆ huge
- ◆ large
- ◆ narrow
- ◆ small
- ◆ thick
- ◆ thin
- ◆ wide

# Navigational Ontology

- ◆ Integrated into the Graphical Ontology.
- ◆ Navigational backbone based on Recursive Navigation Grid.
- ◆ Absolute and relative locations with inference.
- ◆ Location: fuzzy description, points, silhouettes



- ◆ *Family* handling family relationships – useful for family photo albums.
- ◆ *Sights* handling important places of interest.
- ◆ *GoF* handling „Gang of Four“ design patterns – a pilot e-learning application (under construction).

## ◆ Communication modes

- Information retrieval mode
- Image information supplementing mode
- Free communication mode

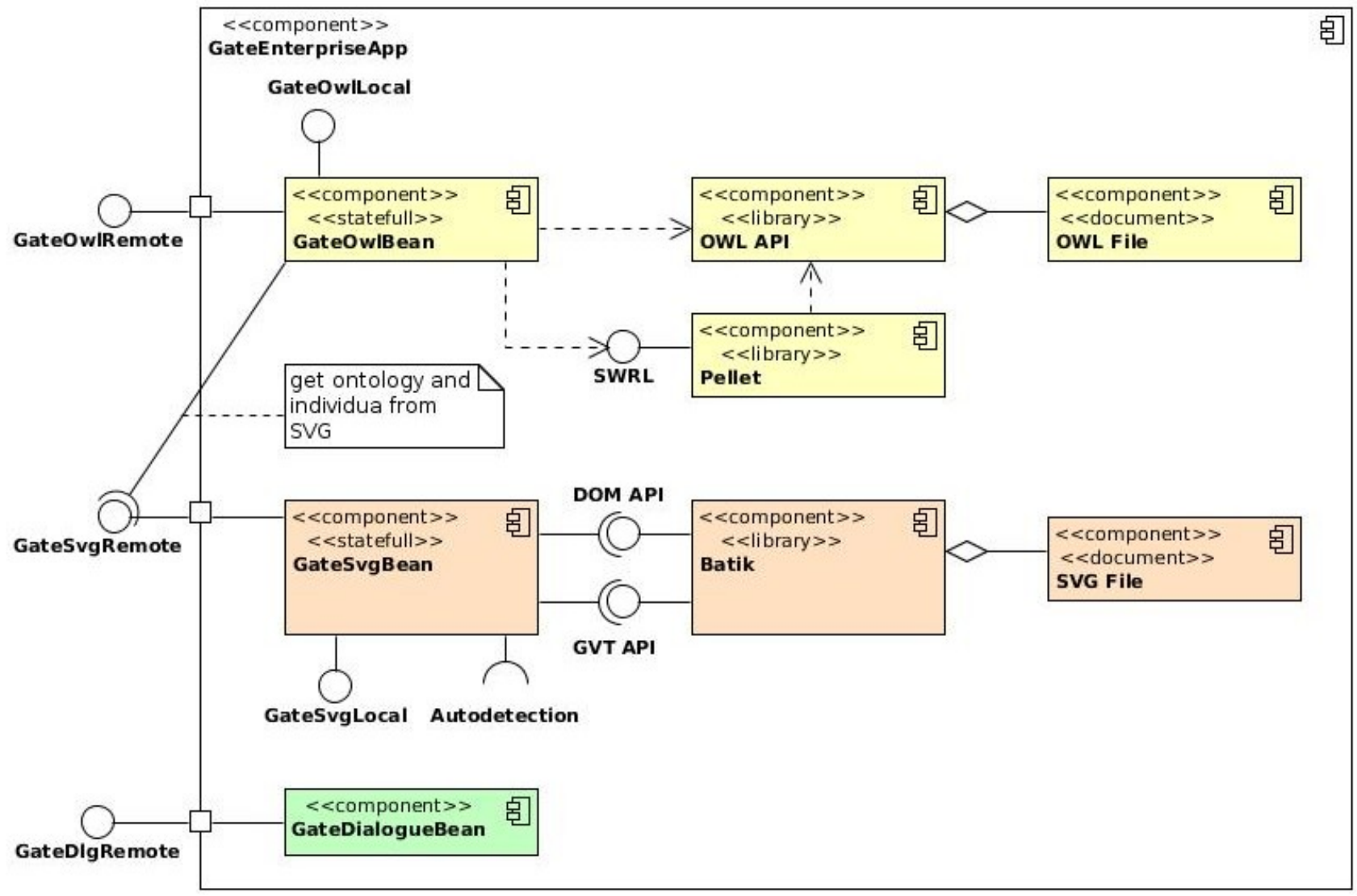
## ◆ Communication analysis

- Domain-specific small fragment of natural language
- Relatively simple grammars
- Frames technology
- Standard techniques for misunderstanding solving
- Example: WWL, What-Where Language

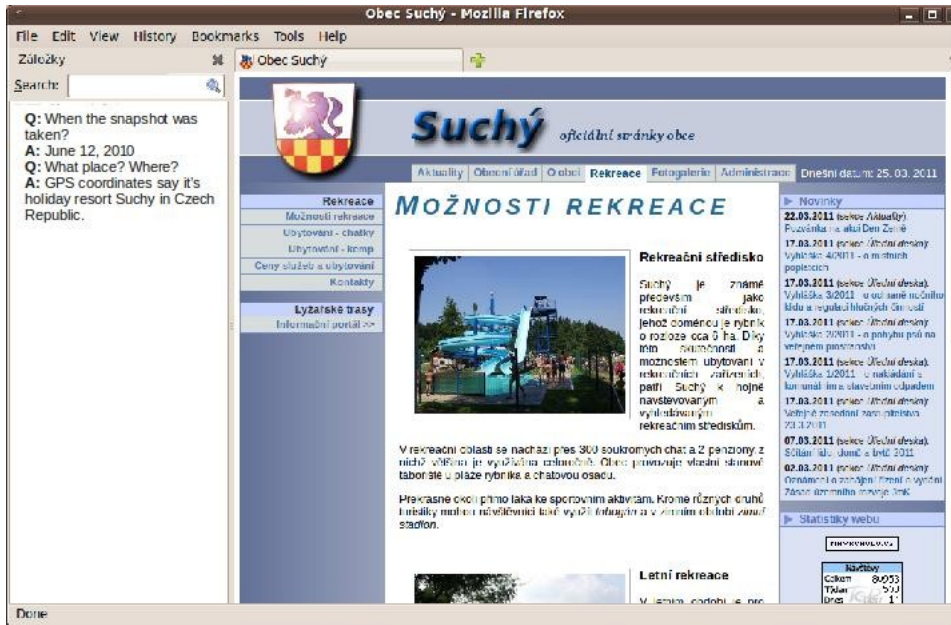
*How far is it from this hotel to the nearest beach?*

*How far is it from <SLOT1> to <SLOT2>?*

# GATE system – server side

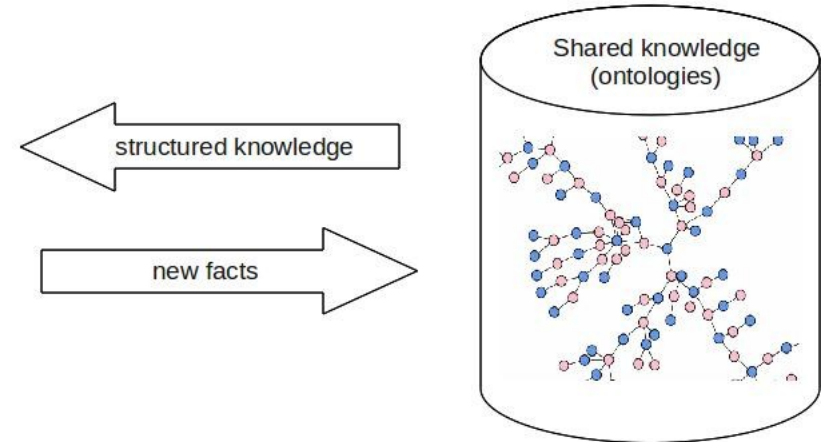


- ◆ The project focuses on web technologies and direct interaction with images on web pages



## Dialogue plug-in to web browsers

- handles initial interaction action, e.g. clicking on a picture
- handles the dialogue window
- communicates with server



## Server side:

- JavaEE, EJB web services
- stores the knowledge base
- applies auto-detection and image recognition algorithms

- ◆ Creating domain-specific ontologies
  - Manually, i.e. for e-learning – laborious and exhausting
  - Dynamically from dialogue – correctness, abstraction
- ◆ Dialogues definition
  - Manually create grammar from ontology and then create frames
  - Automatically generate dialogues from ontologies
    - User's behaviour formally modelled by ontology/logics
- ◆ Information gathering
  - Manually, i.e. semantic data are provided by annotator and they are fixed.
  - Learning from dialogues
    - Direct: „I probably depict mountains. Confirm it, please.“
    - Indirect: e.g. the user question „Who is the lady next to the car“ notifies the image that there is a lady and a car in the image.
- ◆ Getting users involved into the using communication images
  - Specialized application, e.g. e-learning,
  - Integration to social networks, ...

- ◆ Personal photo albums
  - Organize, search
- ◆ E-learning study materials linked to domain-specific knowledge base
- ◆ Applications for people with special need
  - Visually impaired people
  - Older adults

S: In the picture there is the chemical structure of antioxidant resveratrol.

U: What is antioxidant resveratrol?

S: Antioxidant resveratrol consists of two benzene nucleus and three hydroxyl groups.

U: What is benzene nucleus?

S: Benzene nucleus ...

/The system uses definition from the chemical ontology/

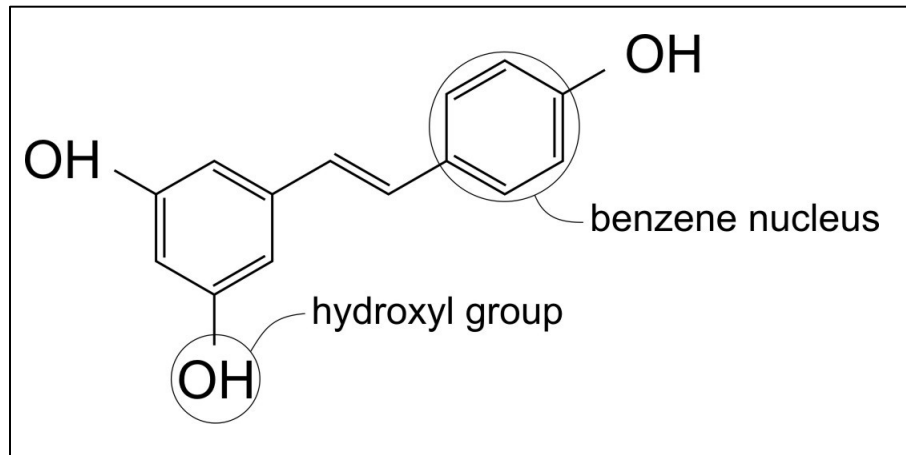
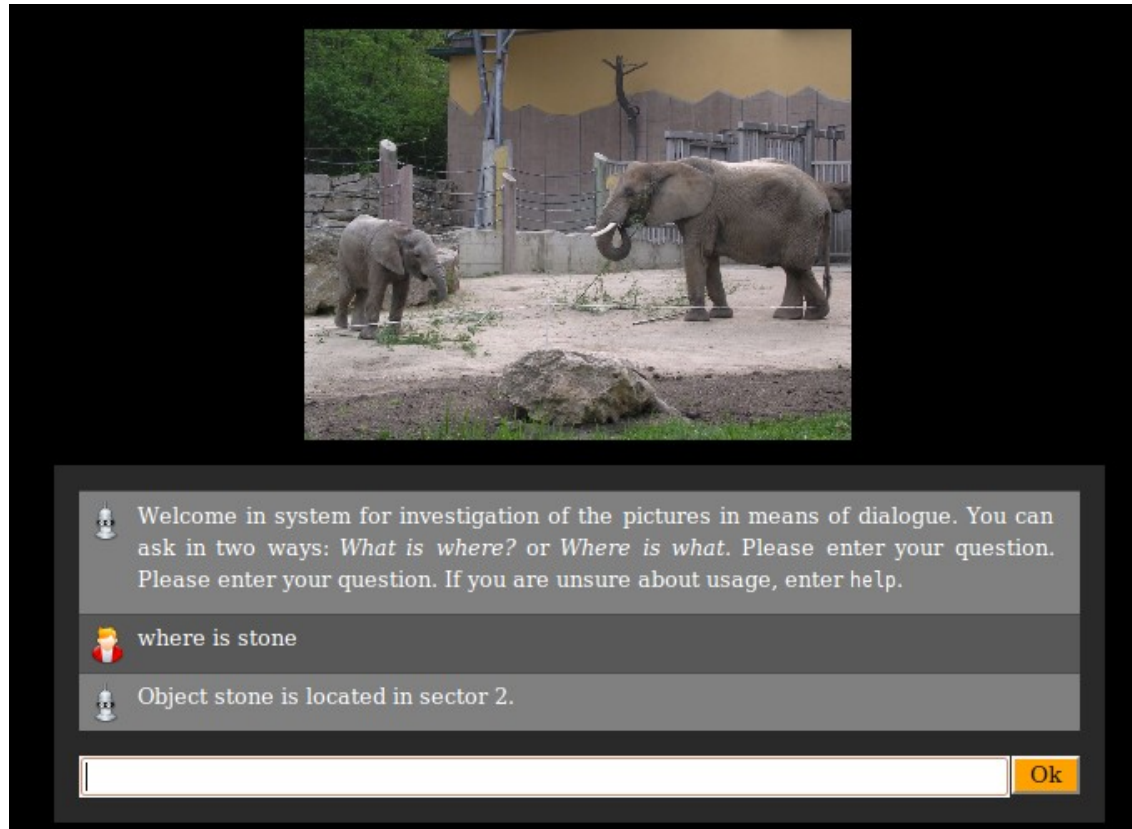


Fig. Antioxidant resveratrol



- ◆ WWL investigation of annotated pictures
  - Web services for the investigation of graphical content by means of What-Where language
  - <http://andromeda.fi.muni.cz/gate/picture-viewer>



where is stone

Object stone is located in sector 2.

## ◆ Painting by dialogue

- Web services for asking objects from database and placing them in desired position of target picture
- <http://andromeda.fi.muni.cz/gate/picture-generator>

U: Put a comet in the sector 9.

U: Put a snowman into the bottom left corner.

U: Write the text „Merry Christmas and Happy New Year“ into the horizontal center, color yellow.

U: Write the text „PF 2010“ into the bottom right corner, color blue.

U: Set background to snowflakes.

U: Generate.



► **Fig.** The Christmas card generated by a blind user

---

Thank you for your  
attention!