画像の意味内容に対するアノテーション仕様書ver.1.2

2013/09/5

# 概要

　本仕様書は、大学入試問題に含まれる画像の表わす意味内容を、自然言語その他の手段によって記述するためのものである。原則として、「理想的な画像認識システムに個々の画像を入力した際の出力」を想定した記述を行う。

# アノテーションの形式

　画像の意味内容の記述を示す範囲は、<adata>タグによってタグ付けする。

<adata> 画像の意味内容の記述

@id 文書内で固有の識別番号。Xに半角数字を続けた形式。（例：X8）

@type 意味内容の記述の表現方法を示す。取りうる値は以下の通り。

text 自然言語のテキスト

table xmlタグによる表の記述

formula 数式

other その他

@field 記述内容を細分化するためのフィールド。主な属性値は以下の通り。

object：画像が表すもの。配置や形状も、特筆すべきことがあれば記述（GoogleImageやTinEyeなどの画像検索にかからなかったものは、極力詳しく記述する方針）。

event：画像が表す出来事や行動。

Google Images： 「最良の推測結果」があれば記述。日本語以外の場合は、日本語訳に原文（カッコ付き）を併記。

TinEye：最上位にヒットしたものの記述を取得。日本語以外の場合は、日本語訳に原文（カッコ付き）を併記。

pointer:地図中の記号と、それが表す地名（現在のもの）。ハイフンでつないで、複数ある場合はカンマで区切る。

caption：図の中に、図の出展などの文字情報がある場合はここへ。

graph：グラフの数値的表現。

label：画像のラベル情報。

comment：その他、伝える必要があることはここへ。

<adata>による記述は、対象となる画像へのリンクを含む<img>タグの直後に配置される。<img>タグのcorresponding\_obj属性は、対応する<adata>のidを参照する（複数の<adata>のidを参照可能）。

* <img>が複数の<adata>に対応している例



<img src=”YozemiCenter-2013--1-SekaishiB-002.jpg” corresponding\_obj=”X2\_1 X2\_2”/>

<adata id=”X2\_1” type="text" field="object">正面に大きな仏像、右側に小さな像。</adata>  
<adata id=”X2\_2” type="text" field="GoogleImages">雲崗石窟（原文：云 冈 石窟）</adata>

<img>タグは、参照用のデータを示す<data>、選択肢群を表す<choices>、個々の選択肢を表す<choice>の直下に現れる場合がある。各ケースにおけるアノテーション例は以下のとおりである。

* <img>タグが<data>の直下にある場合（画像が参照用のデータとして表れている場合）

<data>

<img src=” YozemiCenter-2013--1-SekaishiB-001.png” corresponding\_obj="X1" />

<adata id=”X1” type="....... ">

.....（内容の記述）.....

</adata>

</data>

* <img>タグが<choices>の直下にある場合（選択肢が画像に埋め込まれている場合）

<choices>

<img src=” YozemiCenter-2013--1-SekaishiB-001.png” corresponding\_obj="X1" />

<adata id=”X1” type=".......">

<choice>.....（内容の記述）.....</choice>

<choice>.....（内容の記述）.....</choice>

<choice>.....（内容の記述）.....</choice>

<choice>.....（内容の記述）.....</choice>

</adata>

</choices>

* <img>タグが<choice>の直下にある場合（各選択肢の中に画像が埋め込まれている場合）

<choices>

<choice><cNum>①</cNum>

<img .... corresponding\_obj="X8" />

<adata type="...... " id="X8"> .... （内容の記述）.....</adata>

</choice>

<choice><cNum>②</cNum>

<img .... corresponding\_obj="X9" />

<adata type="...... " id="X9"> .... （内容の記述）.....</adata>

</choice>

<choice><cNum>③</cNum>

<img .... corresponding\_obj="X10" />

<adata type="...... " id="X10"> .... （内容の記述）.....</adata>

</choice>

<choice><cNum>④</cNum>

<img .... corresponding\_obj="X11" />

<adata type="...... " id="X11"> .... （内容の記述）.....</adata>

</choice>

</choices>

# アノテーション内容

写真、絵の場合

* 画像の中に含まれるもの（メインのモチーフ）が何であるかを自然言語で記述する。
  + 写真の場合は、TinEyeやGoogleImages等で検索し、写真から固有名が出てくればそれを記述する。固有名がなければ一般名を記述する。
  + 画像により表現されている状態、行為や出来事があれば記述する。
* 複数の物体が含まれる場合は、画像の中の配置（上下左右）も記述する。
* 出典（画像内にあれば）

地図の場合

* どの地域の地図かを記述。
* 地図内の場所を指し示す記号がある場合は、記号とそこの地名（現在の）のペア（半角ハイフンつなぎ、カンマ区切り）を記述する。
* 出典（画像内にあれば）

グラフ

* 以下のいずれかの最適な方法で記述する
  + 表
  + 数値的表現、数式
  + 概念的な記述（自然言語による記述。縦軸、横軸が何であるか、各曲線が何を示しているか等）
* 出典（画像内にあれば）